Latent factor modelling of disability

Marcello Morciano

A thesis submitted for the Degree of Doctor of Philosophy
in Economics

Institute for Social and Economic Research (ISER)
University of Essex

September, 2015

Summary

This PhD thesis uses survey data and involves the application of latent factor structural equation methods to the study of the economics of disability and disability policy in later life, a topic which is currently very high on the policy agenda.

It comprises four studies. The first chapter investigates the presence of health-related sample attrition (the drop-out of eligible sample members over time) in the English Longitudinal Study of Ageing (ELSA).

The second chapter examines whether different indicators of disability, collected in three widely-used household surveys, are consistent with a common set of findings relating to the targeting of disability benefits.

In the third chapter we estimate the additional personal costs experienced by disabled older people to achieve the same material standard of living as similar people living without disability.

Chapter 4 assesses the presence of socio-economic disparities in birth-cohort trends in later life physical and cognitive disability and in the receipt of non-means-tested cash disability benefits.

Contents

CHAPTER 2:

Do household surveys give a coherent view of disability benefit targeting? A multi-survey latent variable analysis for the older population in Great Britain

CHAPTER 3:

Disability costs and equivalence scales in the older population in Great Britain

Acknowledgments

I express my gratitude to my first supervisor Steve Pudney for his valuable guidance and advice (not only) as a PhD student. I would also like to thank Ruth Hancock for her priceless support during these years at the Health Economics Group at the University of East Anglia.

My appreciation goes to my parents who taught me the value of education, perseverance and dedication. The greatest thanks go to my wife and my son for sustaining me along the way and for filling my days with joy, love and laughter.

Declarations

The present thesis is the result of research I have conducted as a part-time PhD candidate at the Institute for Social and Economic Research (ISER), University of Essex and as a Research Fellow at the Health Economics Group (HEG) at the University of East Anglia.

The research included in this thesis was funded mainly by the Economic & Social Research Council and The Nuffield Foundation.

Chapter 1 is sole-authored. I wish to thank seminar participants for helpful comments and discussions, in particular at the ISER and HEG internal seminars, the Italian Health Economics Association (AIES) conference (University Ca' Foscari, Venice (Italy), October 2014), the Health Surveys User conference 2015 (University College London, London (UK), July 2015) and the seminar at the Department of Statistics, University of Padua (Italy, September 2015).

Chapter 2 has been published in the *Journal of Royal Statistical Society, series A*: *Statistics in Society* (Article first published online: 3 MAR 2015; `DOI: 10.1111/rssa.12107`) and it is a joint work with Ruth Hancock, Stephen Pudney and Francesca Zantomio. An earlier version is available as HEG working paper 13-03: https://www.uea.ac.uk/medicine/health-economics-group/working-papers. S. Pudney conceptualised ideas and supervised all aspects of its implementation. I have derived the dataset from the English Longitudinal Study of Ageing (ELSA), conducted all the statistical analysis, synthesised and interpreted findings. R. Hancock derived the dataset from the Family Resources Survey (FRS) whereas F. Zantomio derived the dataset from the British Household Panel Survey (BHPS). I wrote a substantial part of the papers following my co-authors' useful advice and comments during the development of the work but all authors contributed to the writing of the journal article and reviewing drafts.

Chapter 3 has been published in the *Review of Income and Wealth* (Volume 61, Issue 3, pp. 494–514, September 2015) and it is a joint work with Ruth Hancock and Stephen Pudney. An earlier version is available as ISER working paper 2012-09: https://www.iser.essex.ac.uk/publications/working-papers/iser/2012-09. I have originated the study, conceptualised ideas, synthesised analyses and interpreted findings. R. Hancock derived the dataset from the Family Resources Survey (FRS). R. Hancock and S. Pudney supervised all aspects of its implementation. I wrote a substantial part of the papers following my co-authors' comments during the development of the work. However, all authors contributed to the writing of the article and reviewing drafts.

Chapter 4 is sole-authored. Earlier versions of this study were presented at internal seminars in ISER and HEG and I wish to thank participants for helpful comments and discussions. An earlier version of the chapter, with a simpler statistical approach, has been recently published to *Social Science & Medicine* (July 2015, volumes 136-137, doi:10.1016/j.socscimed.2015.04.035) with the final version written in co-authorship with Ruth Hancock and Stephen Pudney. I have originated the study, conceptualised ideas, synthesised analyses, and interpreted findings, following my co-authors' advice and comments during the development of the work.

Introduction

The world's population has been experiencing significant ageing. Ageing partly results from decreasing mortality, but because chronic and degenerative diseases are more common at older ages, it generally results in an increased prevalence of disability. Disability entails a range of immediate and long-term (financial and psychological) consequences that have important implications for the well-being of the individual, the family, and the society as a whole.

How to promote more effectively healthy and active ageing while building an equitable and financially sustainable welfare system that meets the needs of older adults? This is a key policy challenge for many national and local governments and seeking an answer is no a simple task. To guide this process researchers should make an "*intelligent application of quantitative methods to imperfect data in the hope of illuminating important social issues*" (Cowell, 2000, p. 133). This study reflects my attempt to serve the complex and challenging field of the economics of disability in old age.

This thesis uses survey data and involves the application of latent factor structural equation methods to the study of the economics of disability and disability policy in later life, a topic which is currently very high on the policy agenda. It comprises four chapters which share a similar statistical framework, applied to representative samples of older people living in private dwellings in the United Kingdom. Each chapter is written as a self-contained journal article and thus can be read independently of the others.

The first two chapters deal with the representativeness and completeness of (longitudinal and cross-sectional) survey data, often used for studying disability in old age. Questions, such as 'How ignorable is health-related survey non-response at old age?' and 'Do different measures of disability collected in different surveys provide a consistent picture of the targeting of disability benefits?' are addressed in these chapters, followed by a brief discussion of the policy implications of the results.

The following two chapters provide further results relevant for policy evaluation and policy design of public programmes of support for disabled people. 'How much extra income does a person with a level of disability η and income y need, to be as well-off as (s)he would be with disability at reference level η_0?' is the main research question of Chapter 3 that we aim to address by fitting a structural welfare model and inferring the compensating income variation from estimates of its parameters. 'What are the main determinants of the observed rise in the number of older people receiving disability benefits?'; 'How much of their growth stems from trends in physical and cognitive disabilities?'; 'Is there any diverging trend by socio-economic status?' are the research questions that Chapter 4 aims to answer using historical data to identify the lessons we can learn for the future.

In the following paragraphs we will briefly introduce the measurement issues associated with disability and its consequences as well as the statistical approach that we use in this thesis – latent factor structural equation modelling – followed by a brief overview of the four chapters.

Setting the scene

Establishing a meaningful concept of "disability" is difficult and there is no single agreed definition which suits all purposes (Altman, 2001; Haveman & Wolfe, 2000; WHO, 2002). A medical approach would define disability in terms of deviations from medical norms (e.g., presence of diagnosed conditions). In this thesis we follow a functional approach that focuses on individuals' performance by assessing their ability to perform "normal" tasks and roles by measuring for example, their "functional limitations", "difficulties in performing everyday activities", or restrictions on "activities of daily living".

The "performance" criterion is often used to operationalise the concept of disability in economics (as opposed to medicine) where the purpose is to determine eligibility for public programmes and the need for care services. It is also linked more directly to the sociological concepts of social independence and social functioning than are definitions of disability based on the presence/absence of medical conditions. As we argue in Chapter 2, the functional approach of measuring disability also has analogies with Sen's (1985) concept of "capabilities": disability can be seen as a set of constraints on functioning which originate from the interaction of individuals' personal characteristics (e.g., health impairments broadly defined) with their available goods (assets, income) and the surrounding cultural and socio-economic environment.

Sen's theoretical approach offers several advantages for the current study. First, it is a broad framework able to accommodate the conceptualisation of the disablement process as formalised among gerontologists (see, among others, Johnson & Wolinsky, 1993; Verbrugge & Jette, 1994). It sees damage at the cellular level, eventually influencing functioning at the level of organs, which ultimately restricts the individual's capacity to perform tasks and social roles.[1]

The link with Sen's theory is not only important from the perspective of defining (and then measuring) disability. In Sen's view, disability imposes two types of handicap. It reduces not only the ability to generate an income (earning handicap) but also the ability to convert money into good living (conversion handicap). This framework helps in clarifying one limitation of current income-based analysis

[1] It would accommodate also a definition of disability more closely related to the International Classification of Functioning, Disability and Health, which sees disability resulting from activity limitations and restrictions placed upon participation that emerge from the interaction between functional limitations and an unaccommodating environment (WHO, 2002).

of poverty and inequality that substantially underestimate poverty and inequality for families with disabled members by accounting only for the earning handicap (Kuklys, 2005). As we argue in Chapter 3, the conversion handicap can be taken into account when carrying out analysis of the distributional impact of tax and social security benefit reforms by making some allowance for the additional costs of living that different degrees of disability bring.

How to measure disability?

After a definition of disability has been chosen and a theoretical framework identified, we can tackle the disability-measurement problem. A great deal of attention has been devoted in the literature to developing (a vast range of) measures with satisfactory properties in terms of dimensionality, reliability and validity. The ADL (Katz *et al.*, 1963) and IADL (Lawton & Brody, 1969) are perhaps the most successful examples.

A counting approach (summing the number of functional difficulties reported or "indicators") is typically used to derive a scale measuring the presence and severity of needs. The popularity of this type of aggregation is due to its transparency and ease of interpretation. However, different indicators in a scale may indicate different degrees of disability. For example, needing help to eat is a more severe (and more gender-neutral) indicator of disability than needing help in preparing a meal.

Attempts are sometimes made to scale indicators by “weighting” them on the basis of a theory, experience, or by consultation of experts. In any case, different judgements (e.g. due to a different panel of experts) will result in different value for the weights, which may drive different policy implications. Moreover, if in an available survey two or more scales for disability are collected (i.e. ADL and IADL), is there any valid reason to focus only on one?

Why is a latent factor structural equation approach appropriate in our context?

Two crucial issues in the above measurement approach are the *adequacy* of the chosen indicators in reflecting the corresponding dimension(s) of disability; and the *arbitrariness* in the choice of weights used to combine indicators into a single measurement. We argue that such steps should be transparently justified given the relevance of value judgements in deciding that an indicator is redundant or invalid or that it should count more than another one.

An alternative approach would be to “let the data speak”, with weights assigned to the different indicators not chosen by the researcher, but derived statistically. In this thesis, we follow this alternative and propose the use of a latent factor structural equation approach to the study of the economics of disability in later life.

Latent variable Structural Equation Modelling (SEM) is a general and flexible approach that enables us to test hypotheses about both measurement (how well observed indicators serve as a measurement instrument for the underlying latent construct(s)) and structural relations (how latent construct(s) are related to each other and with a set of observed variables believed to be important determinants or consequences of the latent constructs) simultaneously and within a single framework. Latent factor SEMs, commonly employed in psychology and the social sciences, are becoming increasingly important in economics thanks to the emerging body of research that: establishes the parallel importance of non-cognitive skills (personality, social and emotional traits) as well as cognitive skills in producing social and economic success (Heckman *et al.*, 2006); that operationalises the capabilities approach to welfare economics (Anand *et al.*, 2011) and in the "happiness" literature (van Campen & Iedema, 2007; van Praag *et al.*, 2003).

This approach offers several advantages. First, it assumes that the underlying concept is not directly observable (i.e. is latent) but manifests itself in many observed variables (indicators) (Bollen, 1989).

Second, in contrast with principal component analysis, a latent factor approach imposes a "structure" in the sense that the observed indicators are postulated to be (linear) functions of unobserved latent variables. Using Sen's term, we can see each indicator identifying a particular "achievement" (or the lack of it). Each indicator, however, provides only a partial measure of the underlying (possibly multi-dimensional) individual's disability.

Two main assumptions define the causal mechanisms underlying the responses. First, it is assumed that the responses on the indicators are the result of an individual's position on the latent variable(s). The second assumption, known as the *local independence axiom* (Bollen, 2002), states that there are one or more latent variables that create the association between indicators, and when the latent variables are held constant, the indicators are independent.[2] This is particularly important in addressing possible errors of measurement in the indicators, an issue that – while often neglected in applied research – is very relevant for any application that uses indicators in the form of individual reports of self-assessed measures collected through surveys, especially in the health domain (Bound, 1991).

Recalling that the main idea behind the latent variable approach is that observed indicators are manifestations of latent concepts, it is important to allow the possibility that other exogenous variables might "cause" and influence the

[2] Formally: $P[D_1, D_2, \ldots, D_K\,|\boldsymbol{\eta}] = P[D_1|\boldsymbol{\eta}]P[D_2|\boldsymbol{\eta}] \cdots P[D_K\,|\boldsymbol{\eta}]$ where $D_1, D_2, \ldots, D_K$ are observable indicators of functional disability, $\boldsymbol{\eta}$ is a vector of latent variables, $P[D_1, D_2, \ldots, D_K|\boldsymbol{\eta}\,]$ is the joint probability of the indicators $\boldsymbol{D}$ for given $\boldsymbol{\eta}$, that equals the product of the conditional probabilities, $P[D_1|\boldsymbol{\eta}]P[D_2|\boldsymbol{\eta}] \cdots P[D_K|\boldsymbol{\eta}]$, when the latent variables are responsible for the dependencies among the indicators. A latent factor approach permits the use of both continuous and discrete observed variables as indicators.

latent factor(s). In all chapters we allow latent constructs to vary significantly with observable characteristics. For disability, we emphasise the importance of modelling observable sources of heterogeneity induced by individuals' socio-demographic and economic characteristics (SES). This is not only because of the well-known SES-health gradient (Deaton, 2002; Deaton & Paxson, 2001; Graham, 2009) and the "valuation neglect" problem (Sen, 1985) but also because the relationship between disability and income has important implications for the public cost of disability support policy.

If we assume the presence of two or more latent dimensions, we would expect to model how these mutually influence one another and hence it is important to explicitly specify these interactions in the form of a structural model. This can be done in different ways and this thesis offers a broad range of applications.

Four empirical analyses of the consequences of disability in old age

Longitudinal health-related surveys are valuable sources of data for monitoring population health and for the evaluation and design of policy programmes for disabled people. However, the failure of interviewing all eligible individuals for a survey (unit non-response) can seriously distort results, in particular if the mechanism that causes it is related to the phenomena of interest. In Chapter 1, I investigate the *relevance* and *ignorability* of initial health-related non-response at the beginning of the English Longitudinal Study of Ageing (ELSA). In doing so, I integrate the measurement of latent constructs (one of which is poor health) in a rational choice approach for survey participation. A second latent construct captures the individuals' engagement with the scope of the survey and it is allowed to be correlated with latent poor-health. The "engagement" index, while highly correlated with latent health, plays the most important role in explaining participation so that it dilutes the health gradient with non-response, whose relevance however is not completely eliminated. Structural parameters are then used to derive new survey weights for estimating the distribution of health status and receipt of disability-related benefits among older people interviewed at follow-ups. Results suggest that initial non-response is problematic and mainly determined by additional factors not captured in the ELSA weights.

Chapter 2 compares statistical models of the prevalence of receipt of Attendance Allowance (AA) – the main non-means-tested disability benefit available for older people – estimated from three UK surveys: the British Household Panel Survey (BHPS), the Family Resources Survey (FRS) and the ELSA. It employs a structural equation approach in which probabilities of receiving AA depend on latent disability. Two aspects of this comparability issue are specifically: whether the questionnaire content generates disability indicators that are capable of reflecting all the multiple dimensions of disability (completeness); and whether the different indicators available in the three surveys of any particular dimension of disability give the same undistorted picture of the underlying concept (compatibility). We conclude that compatibility is not a serious difficulty although there are some signs that completeness is a problem for the BHPS.

Chapter 3 applies the compensating variation principle for estimating – parametrically – disability costs among older people. A two-latent factor structural model (which fully recognises the latent nature of the constructs “disability” and “deprivation”) is used to estimate a base-dependent equivalence scale (i.e. one which varies by income level) which takes account of the severity of disability. The estimated costs are large (on average £100 per week, in 2007 prices) and rise significantly with disability. The restrictions on preferences imposed by the assumption of a base-independent equivalence scale for disability are not supported by FRS data, implying that the extra income that disabled people on higher incomes need to be as well-off as their non-disabled counterparts is lower than the equivalent proportion of income needed by disabled people on lower incomes. Comparing the estimated costs of disability with the amounts of existing disability benefits suggests that public provision falls considerably short of total disability costs for older people in Great Britain. Estimates have clear implications also for analyses on the targeting and redistributive efficiencies of existing benefits as well as for the design of disability-related public programmes.

The current and future financial sustainability of public programmes of support for disabled people is at the heart of the recent policy debate. After controlling for the "pure" demographic effect of population ageing, a fundamental question is how much of the growth of social security benefits for disabled people can be explained by trends in the underlying prevalence and severity of disability. In Chapter 4, I use a two-latent factor structural equation approach to estimate the (birth-)cohort-year effects in physical and cognitive functionings and in the receipt of non means-tested cash disability benefits (DBs) for older people born between 1924 and 1945. The chapter also investigates the extent to which the overall disability trends have been more favourable among advantaged than disadvantaged socioeconomic groups and how the public disability programmes reacted. Drawing from a pooled sample of the last 10 years of FRS available, the chapter concludes with a series of relevant messages for current and planned policy reforms aimed at supporting older people with care needs.

A final chapter concludes and sets research plans that build upon these achievements.

References

Altman, B. M. (2001). Disability Definitions, Models, Classification Schemes, and Applications. *Handbook of disability studies*, 97-122.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*: John Wiley & Sons.

Bollen, K. A. (2002). Latent Variables in Psychology and the Social Sciences. *Annual Review of Psychology, 53* (**1**), 605-634.

Bound, J. (1991). Self-Reported Versus Objective Measures of Health in Retirement Models. *Journal of Human Resources, 26* (**1**).

Cowell, F. A. (2000). Chapter 2 Measurement of Inequality. In B. A. Anthony & B. François (Eds.), *Handbook of Income Distribution* (Vol. Volume 1, pp. 87-166): Elsevier.

Deaton, A. (2002). Policy Implications of the Gradient of Health and Wealth. *Health Affairs, 21* (**2**), 13-30.

Deaton, A. S., & Paxson, C. (2001). Mortality, Education, Income, and Inequality among American Cohorts. In D. A. Wise (Ed.), *Themes in the Economics of Aging* (pp. 129-165): University of Chicago Press.

Graham, H. (2009). *Understanding Health Inequalities*: McGraw-Hill International.

Haveman, R., & Wolfe, B. (2000). Chapter 18 the Economics of Disability and Disability Policy. In J. C. Anthony & P. N. Joseph (Eds.), *Handbook of Health Economics* (Vol. Volume 1, Part B, pp. 995-1051): Elsevier.

Johnson, R. J., & Wolinsky, F. D. (1993). The Structure of Health Status among Older Adults: Disease, Disability, Functional Limitation, and Perceived Health. *Journal of Health and Social Behavior*, 105-121.

Katz, S., Ford, A. B., Moskowitz, R. W., Jackson, B. A., & Jaffe, M. W. (1963). Studies of Illness in the Aged: The Index of Adl: A Standardized Measure of Biological and Psychosocial Function. *JAMA, 185* (**12**), 914-919.

Kuklys, W. (2005). A Monetary Approach to Capability Measurement of the Disabled in the Uk *Amartya Sen's Capability Approach* (pp. 75-103). Berlin: Springer Science & Business Media.

Lawton, M., & Brody, E. (1969). Physical Self-Maintenance Scale (Functional Assessment). *Gerontologist, 9*, 179-186.

Sen, A. K. (1985). *Commodities and Capabilities*. New York: North Holland.

Verbrugge, L. M., & Jette, A. M. (1994). The Disablement Process. *Social Science and Medicine, 38* (**1**), 1-14.

WHO. (2002). *Towards a Common Language for Functioning, Disability and Health: Icf*. World Health Organisation.

Chapter 1:

Participation of older people in health-related longitudinal studies. Microlevel evidence at the beginning of the English Longitudinal Study of Ageing (ELSA)*

Abstract:
As with many longitudinal studies, non-response in ELSA is unlikely to be ignorable. We investigate health-related bias in participation in the first wave of ELSA and its consequences for assessing both health and receipt of disability-related benefits among those born in or before 1937 (aged 65+ in 2002) at follow-up. We propose a structural equation framework with two latent factors that builds upon the rational choice approach. Results indicate a non-linear relationship between latent health and future non-response. Controlling for latent survey engagement considerably increases the overall explained variance of the model and significantly reduces the (direct) impact of latent health and other exogenous covariates on survey participation. We find that the selectivity in non-response behaviours has important consequences for making inference from univariate analyses on health status and receipt of disability-related benefits from the subsample of respondents at follow-up. Findings have clear implications for fieldwork procedures.

Keywords: disability, disability benefits, birth-cohort trends, latent factor, structural equation model.

---

* I am grateful for comments and suggestions from Cheti Nicoletti, Amanda Sacker, Ruth Hancock and Stephen Pudney. Data from the ELSA were developed by researchers based at University College London, the Institute for Fiscal Studies and the National Centre for Social Research (NatCen) and are made available through the UKDA. I wish to thank NatCen for helpful clarifications on the nature of the data and for providing geographical indicators used for the analyses. Neither the collectors of the data nor the UKDA bears any responsibility for the analyses or interpretations presented here.

1 Introduction

People older than 65 years of age, and especially those aged over 80 years, are the fastest growing age groups of the population in many countries and also the most demanding of care (Colombo *et al.*, 2011). In such a context, a key policy challenge is to promote healthy and active ageing while building an equitable and financially sustainable welfare system that meets the needs of older adults.

Longitudinal studies such as the US Health Retirement Study (HRS), the English Longitudinal Study of Ageing (ELSA) and the Survey of Health, Ageing and Retirement in Europe (SHARE) provide valuable information to study population health dynamics and associations with economic behaviours and status in the later part of life.

A major challenge for inference with longitudinal studies is that a considerable portion of the designed sample fails to participate, so that many outcomes of interest are confounded with the individual's decision to participate in the study. For instance, if panel members' participation is influenced by their health, then analysis of the socioeconomic determinants and consequences of health based on people who remain in the panel may be biased.

Post-collection weighting adjustment strategies are commonly employed to minimise the selection bias that results from non-response,[3] under the assumption that the mechanism generating non-response is Missing at Random (MAR) (Little & Rubin, 1987); that is, conditional on a set of observables, the non-response mechanism is random.

Although weighting adjustment methods are nowadays highly refined, the main concern remains that information predicting non-response may not be completely observed, leading to violation of the MAR assumption and to invalid inferences about population parameters of interest.

The object of this paper is to investigate the existence of health-related survey non-response among older people and its consequences for assessing their health and receipt of disability-related benefits at follow-up. Our results will help to identify which eligible members are likely to drop out of successive follows-up, and inform both the design of instruments that incentivise retention and post-collection procedures that adjust for non-response.

We begin by presenting our theoretical framework, which builds on the Dillman (1978) and Hill and Willis (2001) rational choice approach for survey participation. The proposed empirical framework makes use – for the first time to our knowledge – of a latent factor structural equation approach.

[3] These typically involve the assignment of weights to sample respondents in order to compensate for systematic differences between their characteristics and those of non-respondents.

We postulate that two latent constructs, together with other exogenous (assumed error-free) variables, influence panel members' participation at follow-up. The first of these is health status, which we assume is imperfectly measured by self-reported indicators of health status and health conditions (see, for example, the original papers of Lee, 1982; Van Vliet & Van Praag, 1987; Wolfe & Behrman, 1984), and correlated with individuals' socio-demographic characteristics (Deaton, 2002; Deaton & Paxson, 2001; Graham, 2009).

Perceived costs and benefits of participation are also assumed to be influenced by an individual's underlying attitude towards the study, measured imperfectly by indicators found highly influential in determining survey participation (Banks *et al.*, 2010; Copas & Farewell, 1998; Hawkes & Plewis, 2006; Hill & Willis, 2001; Kapteyn *et al.*, 2006; Loosveldt *et al.*, 2002; Nicoletti & Peracchi, 2005; Tourangeau *et al.*, 2000; Watson, 2003; Watson & Wooden, 2009; Zabel, 1998). We call this second latent factor "engagement", to emphasise its role in measuring how individuals comply with the scope of the survey. Observed indicators of engagement available for our empirical analysis are: item non-response for questions about financial resources in the previous wave; consent to merge survey responses with administrative data; whether the self-completion questionnaire was returned; and completion of the more intrusive health modules.

The *engagement* index is allowed to covary with individuals' socio-demographic characteristics and to be correlated with latent health, since poor health may limit respondents' ability to retrieve the information needed to answer the most difficult questions (Beatty & Herrmann, 2002) or make them more sensitive to the issue of privacy (Jenkins et al., 2006; Johansson & Klevmarken, 2006).

In the empirical section of this paper, we are concerned with non-response at the beginning of the ELSA study. Non-response in the first wave of a panel is typically much higher and relevant than in subsequent waves (Laurie *et al.*, 1999; Lepkowski & Couper, 2002; Pyy-Martikainen & Rendtel, 2008; Watson & Wooden, 2009) and it is often viewed as the most relevant source of non-sampling errors in a panel. The fact that the ELSA sampling frame (the so-called wave 0) is composed of households that participated in previous Health Survey for England (HSE), enables us to model participation into ELSA wave 1 by using information gathered from previous HSEs.

We find evidence of a non-linear effect of health on older people's retention in ELSA at follow-up, although its significance is considerably reduced when "engagement" is taken into account. Using non-response weights derived from our estimates, we find that non-response bias substantially affects the results of static and dynamic analyses of health status and receipt of disability-related benefits at follow-up.

The paper is organised as follows. We begin by reviewing existing literature on survey participation which is relevant to our empirical application. Section 3 presents the econometric formulation of the general framework. Section 4 summarises the main features of ELSA and discusses sample inclusion at each follow-up, up to wave 6. Section 5 presents estimates of the retention model for the first wave of ELSA, using characteristics observed at wave 0. Estimates of the retention model are then used to build a post-collection adjustment technique based on the Inverse Probability Weighting (IPW) estimator (Horvitz & Thompson, 1952). The performance of the new weights is assessed in describing health status and receipt of disability-related benefits among the subsample of respondents in waves 1 and 2. A final section draws conclusions and suggests directions for future research.

2 Literature review

Many researchers from different disciplines have devoted effort to understand the determinants of survey participation with the aim of finding effective policies to reduce non-response and/or to develop post-collection procedures for dealing with possible sample selection bias.

Sociologists and psychologists are mainly concerned with the psychological and social processes involved in survey participation to trace out good survey practice and develop preventive policies that minimise non-response. Sample representativeness, the role of incentives, and the development of robust post-collection adjustment strategies are issues mainly raised in the statistical and economic literature.

A comprehensive treatment of the literature on survey participation and its determinants is available, for example, in Groves and Couper (1998),Uhrig (2008), Watson and Wooden (2009). By drawing on the sparse literature, this section reviews the salient features relevant to our application.

It has been found that the propensity to participate in a survey varies according to survey design features, the interview situation and interviewer workload (Groves *et al.*, 1992; Groves & Couper, 1998; Hill & Willis, 2001; Lepkowski & Couper, 2002; Nicoletti & Buck, 2004); therefore the well-design of a survey is the key aspect for ex-ante minimisation of non-participation. However, individual characteristics, personality traits and attitudes play important roles in the formulation of the decision to take part in the study (Copas & Farewell, 1998; Groves *et al.*, 1992; Hill & Willis, 2001; Norris, 1985; Tourangeau *et al.*, 2000; Zabel, 1998).

In empirical studies, poor health has often been found to be associated with lower participation in general multi-topic household surveys[4] (Groves & Couper, 1998; Lepkowski & Couper, 2002), but with no common pattern. Empirical studies which specifically explored the retention behaviours of older people in longitudinal health-related survey have typically found mixed results.

More often than not, compared with participants, dropouts have been found to have poor physical health and cognitive impairments (see for example the systematic reviews conducted by Bhamra *et al.*, 2008; Chatfield *et al.*, 2005). Kapteyn *et al.* (2006) found that the onset of health conditions such as diabetes, hypertension, or mental health problems or severe problems such as heart and lung disease or stroke are significantly positively associated with future non-response in the HRS for males but are not statistically significant for women. For women, only the onset of a limitation in performing daily activities has a significant negative effect on survey participation. However, Banks *et al.* (2010) found no significant relation between disease prevalence and future unit non-response in the ELSA and the HRS. Similarly, other studies have found no significant difference in future non-response by self-reported health (Van Beijsterveldt *et al.*,

[4] As an example, Contoyannis *et al.* (2004) found that those reporting very poor initial health had a probability of dropping-out from the British Household Panel (BHPS) between 2 and 6 times greater than those who reported excellent health. In their study however, a dynamic model of health status is estimated correcting for non-response as well as cases of "*individuals becoming ineligible because of incapacity or death*" (p. 288). Because of the strong relation between health status and future death, the inclusion of ineligibles in the category of non-respondent is likely to make the selectiveness of unit non-response stronger than it actually is.

2002) in the Maastricht Aging Study nor by using physical and mental health measures (Deeg *et al.*, 2002) in the Longitudinal Aging Study Amsterdam. Matthews *et al.* (2006) reported that only about 30% of refusals of eligible members for year 6 onwards of the Medical Research Council Cognitive Function and Ageing Study (MRC-CFAS) were in poor health, whereas the remaining 70% were "active" but not willing to be re-interviewed. Deeg *et al.* (2002) reported that some members who refused to participate at follow-up were in better health and reported fewer chronic conditions than participants.

The difficulty in assessing health-related non-response arises largely from the difficulty of measuring individuals' health and we are not aware of any study on survey participation that has addressed all these complexities systematically. The use of just a single health indicator has typically been found to have a large influence on estimated relationships with survey non-participation (Jones *et al.*, 2006; Uhrig, 2008). But health, covers many interrelated dimensions, including the presence or absence of medical conditions (diseases), cognitive and physical functioning, and self-perception of health (Johnson & Wolinsky, 1993; Verbrugge & Jette, 1994). Each of these dimensions may influence costs and benefits of participating in different ways and such relations can also be non-linear. Costs can be very high for cognitively-impaired individuals, and for those in perfect health who may face higher opportunity costs of participating. Benefits might be higher for those with low/mild medical conditions who may benefit more from free medical checks (Tinker *et al.*, 2009) as well as for physically impaired individuals who might be more likely to be at home or may welcome the diversion of an interview (Stoop, 2005).

Socio-demographic and economic characteristics of the eligible sample have been found to be associated with both the probability of being contacted and, once contacted, the probability of co-operating with the scope of the survey. Their estimated relationships, however, vary according to specific features of the survey, timing and also with the set of covariates used in the econometric specification.

Despite the large use of socio-demographics in modelling survey participation,[5] some researchers (see e.g., Zabel (1998)) pointed out their limited impact when indicators of individuals' attitudes, beliefs and perceptions of the scope of the survey are included. Copas and Farewell (1998) proposed to include additional indicators of the interviewers' enthusiasm to respond when modelling survey participation.[6] Hill and Willis (2001) found that the most influential predictors of future non-response in the third wave of HRS were respondents' engagement with the aims of the survey and their cognitive reaction to the questionnaire.

[5] As Watson and Wooden (2009) pointed out, they are the type of variables readily available (to all researchers) for all sample members, at least at the time before the unit non-response is manifested. This would explain their wide use in the empirical analyses of the determinants of survey participation.

[6] Norris (1985) reported that the group of "disinterested" was the major source of dropout at the beginning of a health-related longitudinal study among old residents in Kentucky.

Item non-response and the degree of imputation on crucial variables were also used as proxies of unpleasant experience and low engagement or lack of interest with the scope of the survey. They were all found to be good predictors of future unit non-response (Hawkes & Plewis, 2006; Loosveldt *et al.*, 2002; Nicoletti & Peracchi, 2005; Watson, 2003; Watson & Wooden, 2009; Zabel, 1998).

Individuals' engagement with the scope of the study might be also elicited by indicators of their willingness to comply with additional requests from the survey teams. As an example, Banks *et al.* (2010) found that failure to return the "self-completion questionnaire" was strongly predictive of subsequent non-response in the ELSA. Similarly, Kapteyn *et al.* (2006) found that those who did not consent to data linkage were more likely to drop-out from the HRS study.

As far the econometric approach is concerned, the common procedure is to treat the manifested indicators of low engagement with the scope of the survey as explanatory (proxy) predictors of future choice of remaining in the study. One exception –and in this respect, the study most closely related to our own – was that conducted by Hill and Willis (2001). They adopted a two-stage approach in which indicators were first used to derive factors (assumed to be independent) which then entered as explanatory predictors in the individual's survey participation model.

It is important to note that both the proxy variable and the two-stage approaches are problematic. Specifically, incorporating the observed items as explanatory predictors in the survey-participation model ignores the fact that they "*are highly correlated – so much so that if we include them all as predictors in a model of participation, multicollinearity becomes a serious problem*"(p. 427 Hill & Willis, 2001). Moreover, such proxy indicators – while usually highly correlated – generally contain sizeable measurement errors given their imperfect ability to fully capture the underlying concept intended to measure, so that traditional statistical methods (such as multiple regression, analysis of variance, simultaneous equation) provide biased results (Kline, 2011; Liu, 1988; Wang & Wang, 2012).

A two-stage approach does not take explicit account of the co-variation between latent factors. As an example, Hill and Willis (2001) found in the first step that the latent variable "easy" – a measure of interviewee's cognitive reaction to the questionnaire – loaded positively on ease of remembering and understanding the questions and negatively on the "engagement" with the scope of the survey indicators. In their second step, the engagement index was positively associated with survey participation, whereas contrasting signs were found on "easy" according to whether principal component factors were rotated or not (Table 3, p. 431), thus whether the negative association between the two factors were removed or not.

For our specific focus, it is also important to note that individual health was not included in the set of covariates used by Hill & Willis in modelling survey participation. This would lead to potential miss-specification problems due to the likely correlation of health with the factors “engagement” and “easy”. This would certainly be the case if people with poor health sustain higher costs in retrieving all the information required in formulating a response or are more sensitive to the issues of invasion of privacy (Beatty & Herrmann, 2002; Dunn *et al.*, 2004; Jenkins *et al.*, 2006; Kho *et al.*, 2009).

In the next session we attempt to develop a sufficiently general econometric approach that can take into account these aspects.

3 Empirical framework

In this section, we present an organising framework for our empirical analysis, building on the rational choice approach of survey participation proposed by Dillman (1978) and Hill and Willis (2001). According to this approach, an individual takes part in a survey if the expected utility exceeds the expected costs.

Let S_i^* denote the net utility an individual derives from selecting a certain decision regarding participation in the survey and R_i denote a binary variable indicating his or her actual decision (so $R_i = 1$ if the individual remains in the study and $R_i = 0$ otherwise). Assuming utility maximisation:

$$R_i = 1 \; if \; S_i^* \geq 0, R_i = 0 \; otherwise \tag{1}$$

the net utility S_i^* is determined by evaluating the costs and benefits of participating in the survey.

Non-participation increases with the costs and declines with the perceived benefits of answering. Benefits may be *financial* (such as payments and free gifts), *tangible but not financial* (receiving medical tests and feedback on medical results, the diversion of a home visit for a person) or *psychological* (such as a feeling of taking an active role in society). Costs include the time spent in providing an answer, psychological stress in responding to sensitive questions or in providing an accurate answer, psychological and physical tension in taking part in medical and cognitive tests which may be felt to be too difficult, intrusive or humiliating (Tinker *et al.*, 2009).

The net utility S_i^* can be decomposed into a systematic component, assumed to be a function of observed exogenous variables[7] and a random disturbance term, with a setup that could take a fairly complex formulation, with non-linearities both in variables and parameters.

In our framework, health is viewed as an endogenously determined capital stock (Grossman, 1972), difficult to quantify empirically but proxied by self-reported

[7] The vector of observables would include variables that have been suggested to affect survey participation, available within the data. Since they cannot be generally observed at the time when non-participation is manifested, they are used to be proxied with information gathered from the previous wave, when i was fully respondent.

health indicators. We follow the methodology adopted by, for example, Lee (1982); Van Vliet and Van Praag (1987); Wolfe and Behrman (1984), who interpret self-reported health indicators (general health, anthropometric measures, clinical disease etc.) $\boldsymbol{H}$ as imperfect indicators of underlying health status (h^*).

Latent health is then assumed to vary according to individuals' socio-demographic and economic (SES) characteristics ($\boldsymbol{x}_h$), capturing both observable sources of heterogeneity in "true" health, such as the SES-health gradient (Deaton, 2002; Deaton & Paxson, 2001; Graham, 2009) and differences, for a given level of "true" health, in the individuals' survey reporting style (d'Uva *et al.*, 2011).

Individual engagement with the scope of the survey (e^*) is also treated as a latent concept, imperfectly indicated by a number of observed variables $\boldsymbol{E}$ which covary with $\boldsymbol{x}_e$ to capture possible observed heterogeneity around its mean. The equations for this 2-latent factor structural model are:

$$S_i^* = \boldsymbol{\alpha}_s \boldsymbol{x}_{is} + f(h_i^*; \boldsymbol{\gamma}_s) + \delta e_i^* + \varepsilon_{is} \tag{2}$$

$$\tilde{H}_{ij} = \mathbf{1}(\lambda_{jh} h_i^* + \epsilon_{jh}) \tag{3}$$

$$h_i^* = \boldsymbol{\alpha}_h \boldsymbol{x}_{ih} + \varepsilon_{ih} \tag{4}$$

$$\tilde{E}_{ik} = \mathbf{1}(\lambda_{ke} e_i^* + \epsilon_{ke}) \tag{5}$$

$$e_i^* = \boldsymbol{\alpha}_e \boldsymbol{x}_{ie} + \varepsilon_{ie} \tag{6}$$

Equation (2) gives the net utility S_i^* as a function of: a vector of exogeneous predictors of survey participation $\boldsymbol{x}_{is}$; $f(h_i^*; \boldsymbol{\gamma}_s)$ which represents a flexible continuous function of h_i^*, and e_i^* representing the individual's latent attitude towards the survey; α, γ and δ are the corresponding coefficient vectors and ε_{is} is a random disturbance term.

Equations (3) and (4) represent the measurement and the structural models of latent health, where $j = 1,2, \dots, J$ indices the indicators of latent health; λ_{jh} is the coefficient of the impact of h_i^* on the j-th indicator, ϵ_{jh} is a random measurement error with $E(\epsilon_{jh}) = 0$ and $Cov(\epsilon_{jh}, h^*) = 0$.

Equation (3) is a generalised measurement part which allows the use of dichotomous and ordered categorical indicators in addition to continuous ones. We consider here a situation with $\boldsymbol{H}$ being dichotomous or ordinal (i.e. Likert-scale) indicators which could imply an ordered probit link function generating the observable indicators H_{ij} from its unobservable continuous form $\tilde{H}_{ij}$: $\tilde{H}_{ij} = m$ if and only if $A_{jm-1} \leq \tilde{H}_{ij} \leq A_{jm}$, $m = 1, \dots, M_j$ with M_j being the number of response categories for indicator H_{ij} and A_{jm} are threshold parameters.

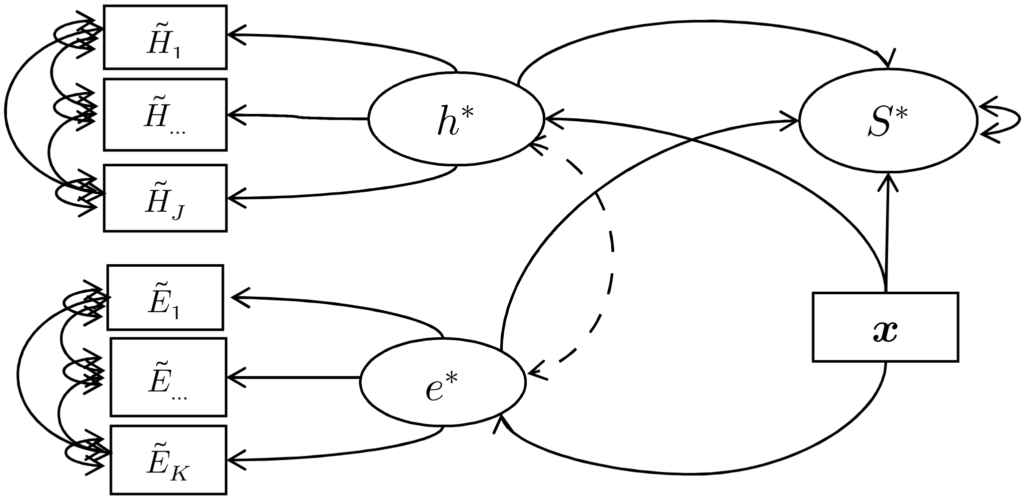
Equations (5) and (6) represent the measurement and the structural components of latent engagement, where $k = 1,2, \dots, K$ indices the observable indicators of latent engagement towards the survey, λ_{ke} is the coefficient of the impact of e_i^* on the k-th indicator, ϵ_{ke} is a random measurement error with $E(\epsilon_{ke}) = 0$ and $Cov(\epsilon_{ke}, e^*) = 0$. The latent variable e_i^* is linearly determined, subjected to a disturbance ε_e, by a set of observable exogenous variables $\boldsymbol{x}_e$ in equation (6), with $E(\varepsilon_e) = 0$ and $Cov(\varepsilon_e, \boldsymbol{x}_e) = 0$.

The model described in equations (2)–(6) is a recursive triangular system of equations for latent health, latent engagement with the scope of the survey and future retention in the study. We allow possible overlaps in the elements of the vectors $\boldsymbol{x_s}$, $\boldsymbol{x_h}$ and $\boldsymbol{x_e}$. We also assume that $Cov(\varepsilon_s, \varepsilon_h) = Cov(\varepsilon_s, \varepsilon_e) = 0$ and that $Cov(\epsilon_{jh}, \epsilon_{ke}) = 0$ and $Cov(\varepsilon_s, \epsilon_{jh}) = Cov(\varepsilon_s, \epsilon_{ke}) = 0$.

Finally, we accommodate the possibility that unhealthy people are less engaged (Beatty & Herrmann, 2002) or have greater sensitivity to issues of privacy (Dunn *et al.*, 2004; Jenkins *et al.*, 2006; Kho *et al.*, 2009) by allowing the error terms for latent health and engagement (equations (4) and (6)) to be correlated.

Figure 1 shows the path diagram for this model. A more general structural model of survey participation with q-latent exogenous variables is developed in Appendix 1.

FIGURE 1: Path diagram of a latent variable structural equation model of the determinant of survey participation



Notes: In path diagrams, latent (unobserved) variables are represented by ovals and observed variables are represented by boxes. Straight one-headed arrows designate direct association. Endogenous variables (indicators) are inter-correlated, as indicated by the bi-directional arrows. Latent variables are correlated as indicated by the bi-directional dashed line.

4 The English Longitudinal Study of Ageing (ELSA)

ELSA is a nationally representative survey collecting data on health and disability and the financial circumstances and well-being of people aged 50 and over and their partners living in private households in England.[8] Many of the health measures adopted in ELSA are comparable with measures used in the HRS (Banks *et al.*, 2010; Wallace & Herzog, 1995) and the SHARE (Börsch-Supan *et al.*, 2005).

The original ELSA sample consists of people born on or before 29 February 1952 and their households, selected from three separate years of the Health Survey for England (HSE) (1998, 1999 and 2001)[9].

In the first wave of ELSA, conducted between March 2002 and May 2003, around 11,500 men and women aged 50 and over and approximately 600 partners aged below 50 were interviewed face-to-face. Surviving sample members living in residential addresses in England were re-contacted every two years, tracking changes in their health and economic circumstances. To encourage response in

---

[8] The ELSA is the result of collaboration between the University College London, the Institute of Fiscal Studies (IFS), and the National Centre for Social Research (NatCen). The universities of Cambridge, Exeter and East Anglia provided expert advice on specific modules. For a fuller description of ELSA we refer to the ELSA user documentation and technical report available at: http://www.elsa-project.ac.uk/.

[9] The HSE is an annual government-funded general health survey of people living in England, carried out by the Department of Epidemiology at University College London and NatCen. Its sample design is drawn from the Royal Mail's small users' Postcode Address Files, stratified by health authority and proportion of households in the non-manual socio-economic groups. User documentation and technical reports are available at http://www.ucl.ac.uk/hssrg/studies/hse.

ELSA, an advance letter was sent to each respondent, giving information about the survey, including the promise of a £10 gift voucher at the end of the interview.[10]

While ELSA was influenced by and modelled on the US-HRS, panel membership in ELSA differs from that of the HRS in one important respect: the ELSA sample is drawn from participants in separate HSE cross-sections. The use of HSE as a sample frame has the advantages of:

a) identifying eligible individuals at reasonable cost;

b) availability of information about respondents' health and other characteristics before they took part in ELSA;

c) increased probability of participation in ELSA due to previous participation in another health-related survey.

The main drawback is that selection into ELSA occurs twice: first as a result of selective non-response in HSE sweeps,[11] and second as a result of selective refusal of HSE sample members to participate in ELSA (Taylor *et al.*, 2007). Non-response and calibration weights were developed by the ELSA team as post-

[10] Average interview length was around one hour and twenty–five minutes for each individual, with high variation by household size and health status. A total of 277 interviewers were used for wave 1. Interviews were clustered and issued to interviewers according to postcode sector. The average number of achieved interviews by interviewer at wave 1 was 44, with a minimum of 2 to a maximum of 112. Given the length of the interview, interviewers were asked to fix an appointment before conducting the interview. The average number of calls to achieve an interview in wave 1 was 3.3, with a minimum of 1 and a maximum of 20. Unfortunately, these details at an individual level are not available to researchers and could not be included in the vector $\boldsymbol{x}_s$ of our study.

[11] Individual response rates in the HSE cross-sections 1998, 1999 and 2001 were 69%, 70% and 67%, respectively (Sholes et. al, 2009; Table 2-1 p. 6).

collection compensation for differences in sampling probabilities, non-response and non-coverage. A detailed description of the ELSA weighting strategy is presented in section 6.

Here we focus on the sample selection problem into the ELSA study. The analysis is restricted to those born on or before 1937 (aged 65+ in 2002), fully interviewed in wave 0, tracking their participation throughout the first six waves of ELSA. This sample selection excludes almost all involvement in the labour force, so that both endogeneity problems of survey participation and labour supply decisions[12] and between health status and labour market participation can be avoided. In addition, over-65s are more likely than younger age groups to experience health conditions and disabilities, so that the presence of health-related non-response is a major concern.

4.1 The unit non-response problem in the ELSA

We distinguish between eligible and non-eligible sample members. Ineligible people are those who died prior to the fieldwork,[13] moved into institutions or outside England, or were erroneously selected as part of the sampling frame (Taylor *et al.*, 2007). Eligible people are divided into respondents and non-respondents.

[12] For those people in paid work, in fact, the opportunity cost of participating in the survey would be rather different than for a retired counterpart.

[13] A linkage with the national registration system enables the identification of deceased people before attempting an interview.

Respondents are those who completed interviews in person (*full*), who were interviewed by proxy or were partially interviewed in person (*partial*). The non-respondent group contains those who were non-contactable or who were contacted but refused to co-operate.

Table 1 shows the outcomes of interviews for the designated sample of members born in or before 1937. From the 9,840 respondents selected from wave 0, 1,005 were ineligible for wave 1: about 96% of those who became ineligible died before the first ELSA interviews took place. The high proportion of deaths observed from wave 0 to wave 1 appears to be connected to the choice of the sample frame which reflects the household composition at the time when the HSE interview took place. From that time to the first ELSA interview (wave 1), between one and four years elapsed, with the result that most of the oldest people selected for the ELSA study died.

The sample contains 8,835 eligible adults aged 65 and over for wave 1. Of these, about 63% were interviewed in wave 1 with 96% providing a full interview, 31% refused and 6% were not contacted. Between waves 1 and 2, 670 sample members became ineligible. A significant subgroup of eligible sample members was not issued for follow-up in wave 2 because *all* members of their household had explicitly refused to be re-contacted after wave 1. For reasons still under investigation,[14]

[14] Personal communication with the ELSA team.

Table 1: Sample frame, non-eligible and eligible members in the ELSA

Wave (fieldwork period)	INELIGIBLE	ELIGIBLE							
	Total (cumulative)	Respondents				Non respondent			
		Full main interview	Partial	**Total**	*As % of eligibles*	contact failure	refusal	**Total**	*As % of eligibles*
0 (1998, 1999 and 2001)*	-	9,840		**9,840**	-	-	-	-	-
1 (3/2002 - 3/2003)	**1,005**	5,371	223	**5,594**	*63.32%*	497	2,744	**3,241**	*36.68%*
2 (6/2004 - 7/2005)	**1,675+ 2,941[a]**	4,055	99	**4,154**	*79.52%*	270	800	**1,070**	*20.48%*
3 (5/2006 - 8/2007)	**5,173**	3,307	164	**3,471**	*74.37%*	161	1035	**1,196**	*25.63%*
4 (5/2008 - 7/2009)	**5,842**	2,730	184	**2,914**	*72.89%*	165	919	**1,084**	*27.11%*
5 (6/2010 - 7/2011)	**6,260**	2,361	214	**2,575**	*71.93%*	172	833	**1,005**	*28.07%*
6 (5/2012 - 6/2013)									

Note: (*) Wave 0 refers to the selected sample born in or before 1937 (aged 65+ in 2002) who provided a full interview in wave 0. Our analysis uses the file `index_file_wave_0-wave_5_v2` provided by ELSA team (last access 13/10/2014), mainly using the wave specific variables `outind` and `issue`. Outcomes of the interview are not currently available for wave 6. **Contact failure** is defined as cases of no contact, broken appointment, away/ill in hospital during survey period, physically/mentally unable/incompetent, contact made but not with eligible resident, address not attempted, address inaccessible, unable to locate address, duplicate address, moved – unable to trace". **Refusal** is defined as cases of refusal before or during the interview, ill at home during survey period, productive interview but respondent requested deletion. See text for details. ([a]) 2,941 eligible members were not issued for follow-up in wave 2 because all wave 1 respondents in the household explicitly refused to be re-contacted and were consequently considered ineligible by the ELSA team. See text for details.

they were flagged as ineligible members by the ELSA team (Scholes *et al.*, 2008 p.40). In our designated sample they number 2,941 and for our analysis in section 6.2 they are treated as cases of non-response. Whether they are treated as ineligible or non-responders makes a substantial difference. For example, the response rate for those born by 1937 at wave 2 is 80% if they are treated as ineligible, and 51% if they are treated as non-respondents.

A number of important messages emerge from the numbers presented in Table 1. Firstly, only about a quarter of the initial sample of those born in or after 1937 can be tracked to wave 5. This is not surprising given the age of the sample and the associated risk of becoming ineligible.

Among eligible members, there was a heavy loss at the beginning of the study. This is in line with reported lower response rates in the first follow-up of other longitudinal studies (Fitzgerald *et al.*, 1998; Laurie *et al.*, 1999; Pyy-Martikainen & Rendtel, 2008; Watson & Wooden, 2009).[15] High dropout rate at the beginning of the ELSA study motivates the focus of this study.

Second, we are able to distinguish ineligibles from non-respondents. The distinction between non-participation due to ineligibility and unit non-response is of paramount importance when we wish to make inferences that related to the

[15] Lepkowski and Couper (2002) argue that the response process in the first wave is fundamentally different from that of subsequent waves. This is both because of self-selection of the least committed sample units and because of the extra information and organisational experience gained by the survey agencies at each follow-up.

whole population. As pointed out by Nicoletti and Peracchi (2005), dropouts because of ineligibility reproduce the dynamics of the target population whereas survey non-response truncates the sample incidentally. For our sample in ELSA, ineligibility is mainly due to death, as the link with mortality register available for ELSA has documented.

Finally, non-response in ELSA is much more often due to refusal than to contact failures. The low rate of contact failure could reflect ELSA's specific focus on older people, who have relatively low household mobility, and its sample being selected from those already participating in another survey. Given this, we combine non-contacts and refusals in a single class of non-respondents for our analysis.[16]

5 Results

The model comprising equations (2)–(6) has been estimated simultaneously allowing for the discrete nature of the dependent variables, using robust maximum likelihood as implemented in *MPlus* version 7.3 (Muthén & Muthén, 1998-2012). Standard errors were clustered by household to allow for intra-household

[16] As an example, Nicoletti and Peracchi (2005) allow for conditional correlation in a bivariate probit model which distinguishes between contact and co-operation in the European Community Household Panel. They found that estimated coefficients were almost invariant when conditional independence was relaxed.

correlation. The final outcome (R) in the subsequent wave is categorised as follows: 1 fully or partial respondent; 0 non-contact, refusal or interview by proxy[17]. The sample is restricted to those born in or before 1937 (aged 65+ in 2002), fully interviewed in wave 0 and eligible for wave 1. Excluding eligible sample members with missing values in our list of explanatory variables measured at baseline, the sample size reduces from 8,835 to 8,420, with a loss of 415 observations.

5.1 Descriptive statistics

A complete list of variables and their descriptive statistics by final outcome of the interview in wave 1 used for the empirical part of this paper is given in Table 2. Women are more prone to drop out from the study, in line with the idea that older women, although more readily contacted, tend to be more sensitive to issues of invasion of privacy (Johansson & Klevmarken, 2006) and less prone to participate in health-related studies (Lynn & Clarke, 2002). Also in line with literature, unit non-response is higher among older people (Lepkowski & Couper, 2002; Uhrig, 2008), un-partnered (Groves & Couper, 1998), non-white (Zabel, 1998), those with low education and those in lower social classes (Groves & Couper, 1998). Being a homeowner (Lepkowski & Couper, 2002) and living in a non-urban

[17] Proxy interviews in subsequent waves are coded as non-respondent because self-reported and objective information on individual health status/disability was not collected for proxy respondents. Thus, proxy interviews are generally excluded in studies on health and health-care utilisation.

area (DeMaio, 1980; Stoop, 2005) is positively associated with survey participation.

Two types of health indicator are consistently collected in the three waves of HSE from which ELSA wave 0 sample was drawn.[18] One is the individual's self-assessed health (SAH) as very bad, bad, fair, good or very good. Those with bad and very bad self-assessed health are more likely to drop out from the study, in line with previous studies. Conversely, those in very good health are more likely to take part in wave 1.

The second type of indicator is related to self-reported medical conditions. In line with Banks *et al.* (2010), descriptive analysis suggests that once ineligibles are excluded from the analysis, the group of non-respondents in wave 1 does not report significantly higher prevalence of medical conditions than respondents. The only exception is for the indicator of cardiovascular problems, which suggests a counter-intuitive positive relationship between that condition and future response.

Engagement with the scope of the survey is significantly higher among respondents in wave 1, providing an early and consistent indication of the importance of such indicators in explaining future response in the ELSA. Dramatic differences

[18] Strictly speaking, the HSE uses modules of questions on specific issues that vary year on year and our empirical analysis uses indicators consistently collected in the three HSEs from which the initial ELSA sample was drawn. It should be noted, however, that the econometric framework proposed can be extended by allowing *missing by design* of some of the endogenous indicators.

Table 2: Descriptive statistics ELSA wave 0 sample according interview outcome in wave 1

	Non-respondent		Respondent		Difference
	mean	*sd*	mean	*sd*	
Covariates					
Women	0.579	*0.494*	0.543	*0.498*	0.0364**
Age of the respondent at year 2002[a]	75.302	*6.882*	74.069	*6.478*	1.2320***
Have a partner	0.588	*0.492*	0.610	*0.488*	-0.0220*
Left education before 14yrs old	0.537	*0.499*	0.462	*0.499*	0.0749***
Left education after 19yrs old	0.056	*0.230*	0.077	*0.266*	-0.0210***
Non white	0.029	*0.168*	0.021	*0.143*	0.0080*
Home owner	0.729	*0.444*	0.762	*0.426*	-0.0330***
Social class: manual worker	0.519	*0.500*	0.462	*0.499*	0.0570***
Interviewed in 1998	0.361	*0.480*	0.433	*0.496*	-0.0722***
Interviewed in 1999	0.196	*0.397*	0.190	*0.392*	0.0058
Interviewed in 2001	0.443	*0.497*	0.377	*0.485*	0.0664***
Urban >10k	0.788	*0.409*	0.760	*0.427*	0.0279**
Health indicators					
Self-assessed health (SAH):					
Very bad	0.040	*0.195*	0.021	*0.145*	0.0181***
Bad	0.091	*0.288*	0.076	*0.265*	0.0150*
Fair	0.301	*0.459*	0.284	*0.451*	0.0174
Good	0.359	*0.480*	0.378	*0.485*	-0.0192
Very good	0.209	*0.407*	0.241	*0.428*	-0.0313***
Medical conditions (suffer from):					
Infectious disease	0.033	*0.177*	0.029	*0.168*	0.0033
Neoplasms & benign growths	0.098	*0.297*	0.103	*0.304*	-0.0055
Endocrine & metabolic	0.022	*0.148*	0.019	*0.138*	0.0028
Blood & related organs	0.035	*0.184*	0.034	*0.182*	0.0008
Mental disorders	0.062	*0.241*	0.059	*0.236*	0.0028
Nervous system	0.049	*0.216*	0.053	*0.225*	-0.0041
Eye complaints	0.275	*0.447*	0.274	*0.446*	0.0016
Ear complaints	0.099	*0.299*	0.100	*0.300*	-0.0009
Heart & circulatory system	0.075	*0.263*	0.091	*0.288*	-0.0163**
Respiratory system	0.033	*0.178*	0.033	*0.178*	0.0001
Digestive system	0.019	*0.135*	0.016	*0.125*	0.0027
Genital-urinary system	0.333	*0.471*	0.332	*0.471*	0.0005
Musculoskeletal system	0.009	*0.096*	0.009	*0.093*	0.0006
Engagement indicators					
Consent link survey data with administrative data	0.806	*0.395*	0.958	*0.201*	-0.1520***
Complete and return the self-completion booklet	0.913	*0.282*	0.972	*0.164*	-0.0592***
Non-missing value at financial questions	0.958	*0.200*	0.978	*0.147*	-0.0196***
Consent having a nurse visit	0.793	*0.405*	0.958	*0.201*	-0.1650***
Observations	3,012		5,408		

Sample size: 8,420 see text for details. *Notes:* * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. ([a]) collapsed at 90.

in future interview outcomes are found in the percentage of those who consent to link survey records with administrative records, those who did complete/return the self-booked questionnaire, those who gave answers to the income questions in the HSE and those who consented to a nurse visit.

5.2 The measurement equations

Table 3 presents results from the measurement equations in (3) and (5) of the full SEM comprising equations (2)–(6).

Factor loadings $(\lambda_{jh}, \lambda_{ke})$, which represent respectively the effects of latent health (h^*) and latent engagement (e^*) on manifested indicators $\boldsymbol{H}$ and $\boldsymbol{E}$, are in line with expectation. The factor loadings associated with latent heath index are positive and highly significant for all the indicators in use. This means that $\boldsymbol{h}^*$ is a decreasing function of health. We also report the squared correlation of each indicator with the underlying latent construct.

It is important to note that we combine items from different health concepts to extract a single latent factor of poor-health,[19] without assuming any temporal/causal ordering. A broad accepted conceptual framework (see, among others, Johnson & Wolinsky, 1993; Verbrugge & Jette, 1994)

[19] Exploratory principal component factor analysis on health-related items indicates the existence of a single latent factor with an eigenvalue of about 2. The second factor has an eigenvalue of 1.1, which is very weak and does not strictly fall below the conventional 1.0 cut-off. Thus, a unique latent poor-health factor, explaining about 20% of the variance in the 14 health-related items, is assumed in the econometric specification.

Table 3: Factor loadings and squared correlations of observables indicators with latent indices

indicator	Factor loadings	*S.E*	R^2
Indicators of h^*			
SAH - Self-assessed health status (1 very good;...; 5 very bad)	4.458***	0.257	0.960
(D) Infectious disease	0.313***	0.031	0.100
(D) Neoplasms & benign growths [a]	0.291***	0.020	0.086
(D) Endocrine & metabolic	0.316***	0.032	0.099
(D) Blood & related organs	0.245***	0.028	0.064
(D) Mental disorders	0.212***	0.024	0.045
(D) Nervous system	0.135***	0.024	0.019
(D) Eye complaints	0.459***	0.018	0.189
(D) Ear complaints	0.398***	0.022	0.150
(D) Heart & circulatory system	0.334***	0.022	0.110
(D) Respiratory system	0.250***	0.030	0.065
(D) Digestive system	0.211***	0.037	0.047
(D) Genito-urinary system	0.459***	0.017	0.187
(D) Musculoskeletal system	0.221***	0.043	0.049
Indicators of e^*			
(D) Consent link survey data with admin data	1.126***	0.102	0.739
(D) Complete and return the self-completion booklet	0.538***	0.039	0.399
(D) Respondent provides a non-missing value at financial questions (income sources and savings)	0.205***	0.030	0.063
(D) Consent having a nurse visit	0.873***	0.063	0.659

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. ([a]) we excluded minor skin cancers. (D) denotes dummy variables. R^2 is the squared correlation between factor loadings and latent factors. Estimates in the table refers to specification “D” (see later in the text) but factor loadings obtained from other specifications are virtually identical.

would see damage at the cellular level eventually influencing functioning at the level of organs, which ultimately restricts the individual’s capacity to perform tasks and social roles. That is, disease leads to impairment, which leads to functional limitations, which in turn cause disability and affect SAH. Therefore one can view medical condition outcomes as *distal* and SAH as *proximal* measures of current latent poor-health.

In our application, we treat all $\boldsymbol{H}$ as being determined by h^* for mainly two reasons. First, causal or temporal ordering cannot be conclusively assessed in cross-sectional data as in our case. Second, a given disablement process can prompt new pathologies and associated dysfunctions, leading to endogeneity issues.[20] However, we found that SAH has substantially higher squared correlation with h^* than the others, though we cannot know for sure that this is due to the higher explanatory power of SAH with respect to medical condition indicators or because it is a 5-point scale indicator with a higher statistical explanation power than binary ones.

The factor loadings associated with latent engagement are all positive and highly significant.[21] The highest correlation with the engagement index is found for the indicators of consent to merge survey data with social security records and participation in the nurse visit. The indicator most widely used by other researchers, item non-response on income questions, is the least sensitive, with

[20] For example, a woman with painful arthritis may reduce her mobility which eventually has an impact on her SAH and starts a vicious circle with possible reduction of her cardiopulmonary function and ability to participate in social activities, with potential effects on her mental health. While detecting the chronological/causal path would be an interesting line of research, our attempt here is limited to describing a broader concept of health which goes beyond the one-indicator – and almost atheoretical – vision of "true" health.

[21] Exploratory principal component factor analysis suggests the presence of a single factor of engagement given that the eigenvalue associated with a second factor falls far below the conventional 1.0 cut-off.

almost 94% of its variance unassociated with latent engagement.

5.3 The structural equations

Structural parameters of the health and engagement equations (4) and (6) are provided in Table 4. The latent poor-health index (h^*) is higher for non-white women and is non-linearly associated with age of the respondent.[22] SES-related differences in health are clearly evident by looking at the magnitude and significance of the coefficients associated with educational attainments, social class and home-ownership. Highly educated individuals have significantly better average health than otherwise similar less educated people. Similarly, high social class and housing wealth are also associated with better health.

TABLE 4: Structural parameters of the health (h^*) and engagement (e^*) equations

Covariate	h^*		e^*	
	Coefficient	*S.E*	*Coefficient*	*S.E*
Female	0.050**	0.024	-0.260***	0.045
Spline age 65-74	-0.006*	0.001	-0.016	0.011
Spline from age 74+ [a]	0.017***	0.003	-0.042***	0.008
Married/cohabiting	0.032	0.028	0.149**	0.066
Completed education before 14 years old	0.186***	0.028	-0.055	0.066
Completed education after 19 years old	-0.245***	0.051	-0.172	0.117
Non-white	0.543***	0.083	-0.865***	0.154
Home-owner	-0.302***	0.031	0.137*	0.072
Social class: manual worker	0.177***	0.027	-0.187***	0.061

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. ([a]) collapsed at 90. Estimates in the table refers to model D specification (see later in the text). The model also includes dummy variables on region of residence, if living in urban area (>10K), and from which HSE cross-section the eligible person was drawn. Standard errors were clustered at household level.

[22] Sample members' age enters in all structural equations in the form of a spline with a knot at the median age (74). It accounts for possible non-linearities in its relation with health (equation (4)), engagement (equation (6)) and retention to the study (equation (2)).

Engagement with the scope of the survey (e^*) is lower for older non-white women and higher for partnered individuals. One might expect a SES-engagement gradient, in line with the idea that high SES individuals are more likely to perceive the social benefits of complying with the scope of the survey (Uhrig, 2008), but our results are not consistent with that. Home-owners show a higher level of engagement (significant at 10% level) and low social class is negatively associated with engagement (p-value <0.001) but, surprisingly, we do not find any significant association of $\boldsymbol{e}^*$ with level of education. This result is consistent with findings in Jenkins *et al.* (2006) where education was not significantly associated with consent behaviours to data linkage.

Table 5: Estimates of retention in the survey in wave 1 of ELSA

Covariate	Model A		Model B		Model C		Model D	
	Coefficient	S.E	Coefficient	S.E	Coefficient	S.E	Coefficient	S.E
Female	-0.107***	0.023	-0.111***	0.023	0.001	0.032	-0.002	0.032
Spline age 65-74	0.001	0.007	0.001	0.007	0.011	0.009	0.010	0.008
Spline from age 74+ [a]	-0.027***	0.004	-0.027***	0.004	-0.012**	0.006	-0.012**	0.006
Married/cohabiting	-0.079**	0.036	-0.081**	0.036	-0.175***	0.048	-0.177***	0.048
Completed education before 14 years old	-0.096***	0.036	-0.095***	0.036	-0.099**	0.048	-0.097**	0.048
Completed education after 19 years old	0.082	0.064	0.087	0.065	0.208**	0.082	0.211**	0.082
Non-white	-0.230**	0.102	-0.226**	0.102	0.155	0.137	0.157	0.137
Home-owner	-0.001	0.041	-0.003	0.041	-0.061	0.052	-0.063	0.052
Social class: manual worker	-0.104***	0.033	-0.104***	0.033	-0.039	0.044	-0.040	0.044
h^*	-0.042***	0.016	-0.047***	0.016	-0.016	0.021	-0.023	0.021
$(h^*)^2$			-0.030**	0.013			-0.026*	0.015
e^*					0.521***	0.030	0.520***	0.03
$cov(h^*, e^*)$					-0.072**	0.030	-0.073**	0.03
Free parameters	45		46		55		56	
Log-likelihood	-49783.925		-40620.272		-49434.928		-49433.378	
Correction for non-normality factor	1.1286		1.548		1.1333		1.1321	
AIC	99769.85		81452.544		99075.856		99074.756	
BIC	100480.725		82198.611		99800.808		99806.746	

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. ([a]) collapsed at 90. All models also include dummy variables on region of residence, if living in urban area (>10K), and from which HSE cross-section the eligible person was drawn. Standard errors were clustered at household level.

Structural parameters of the retention model (equation (2)) are provided in Table 5 in four different variants. *Model A* is a reduced version in which latent health enters linearly and we do not control for latent engagement. *Model B* introduces a quadratic term for health aiming at testing its possible non-linear relationship with retention in wave 1. *Models C* and *D* introduce latent engagement in the set of covariates when h^* is entered linearly and in a quadratic form.

For model A, we found results in line with previous research. Older married women are less likely to remain in the study as well as the non-white population. Lower education and social class are negatively associated with participation in wave 1 of ELSA but we did not find any significant relationship with home ownership.

In model A, there is a positive relationship between health and survey participation: poor-health significantly reduces the likelihood of retention in wave 1 but model B, which fits the data slightly better than model A, reveals a significant U-shape relationship, meaning that, while unhealthy individuals are less likely to remain in the study at follow-up, very healthy individuals do not show significantly lower dropout probabilities with respect to the remaining sample population. This is consistent with the view that healthy people might have concerns about the time cost of participating and in taking part in medical and cognitive tests which, apart from being time-consuming, may also be felt to be humiliating.

Models C and D introduce the latent engagement index in the set of covariates for retention when latent health enters in the model linearly and in a quadratic form, respectively. Both specifications show that the latent engagement index plays the most important role in explaining retention decision for wave 1, by increasing the explained variance significantly, as shown from the goodness of fit statistics available at the bottom of Table 5.

Controlling for engagement, the effect of other covariates is significantly weakened. Gender, race, home-ownership and social class no longer play a significant effect in explaining retention in the study. The coefficient associated with the second spline of age[23] is almost halved, whereas the effect of being partnered is now about 2.2 times higher those obtained in models A and B. Similarly, the effect of level of education is significantly increased, with an emerging significant retention bias in favour of those more educated.

Controlling for engagement, latent health is no longer significantly related to wave 1 participation when its effect is assumed to be linear (model C). The coefficient of the quadratic term of h^* in model D is now significant only at the 10% level. Both specifications provide evidence that, controlling for the individual's engagement at baseline, health plays a minor role in explaining participation in wave 1.

There is a small but significant correlation between h^* and e^* of about -0.07,

[23] See previous footnote for a definition.

p<0.05, consistent with the idea that people in poor health face higher costs in retrieving all the information required in formulating a response or are more sensitive to privacy concerns (Beatty & Herrmann, 2002; Dunn *et al.*, 2004; Jenkins *et al.*, 2006; Kho *et al.*, 2009) and are less engaged. This has to be taken into account when drawing conclusions from estimates in Table 5. A series of robustness checks are confined in the Appendix. For example, constraining the correlation between the two latent constructs to zero increases the importance of h^* in explaining retention in wave 1 but leaves other structural coefficients virtually unaffected (see Appendix Table A1).

5.4 Implications

Figure 2 shows the implications of the estimates of model D in table 5, for three illustrative 75-year old individuals, each with a low level of education and living alone.[24] Engagement with the scope of the survey is set at the values observed at the 25th, 50th and 75th percentile of its Empirical Bayes' (EB) sample prediction[25] distribution. Predicted probabilities of remaining in the study in wave 1 vary considerably according to the level of engagement at wave 0. The weakly engaged individual has on average a predicted retention probability of

[24] Given the non-significance of other $\boldsymbol{x}_s$ but age, marital status and education in explaining retention in wave 1, we set their values to zeros.

[25] EB predictors of the latent variables $\boldsymbol{\xi}\{h^*, e^*\}$ are the means of the empirical posterior distribution with the parameter estimates $\boldsymbol{\theta}_{(.)} = \{\boldsymbol{\lambda}_{(.)}, \boldsymbol{\alpha}_{(.)}\}$ replaced with their estimated model parameters $\widehat{\boldsymbol{\theta}_{(.)}}$ and are calculated by approximation of the following multivariate integral: $\int \boldsymbol{\xi}\omega\big(u|\boldsymbol{H}, \boldsymbol{E}, \boldsymbol{x}; \hat{\boldsymbol{\theta}}\big)d\boldsymbol{\xi}$.

about 38%; rising to 54% and 65% at the medium and high levels of engagement.

Retention rates increase with latent health up to a certain level and decrease slightly thereafter. For a low (high) engaged individual, the predicted probability of participating in wave 1 of ELSA is about 38% (64%) in the bottom 5% percent of the latent poor-health distribution, 40% (66%) at its median level, felling to 34% (60%) for those in the top 5% of the latent health distribution.

FIGURE 2: Predictions of the probability of remaining in the ELSA study by latent health index for three 75-year-old representative individuals with different level of engagement with the scope of the study

Notes: Predicted probabilities computed using estimates of model D in Table 5 (see text for details). To facilitate the interpretation, the labels for the x-axis in the figure refer to the categories of the 5-scale SAH indicator (the most influential factor in determining latent health) which corresponds to the mean values of the EB prediction of latent health observed in the 5 categories of the SAH question.

6 Post-estimation analysis

In this section we assess whether the sample comprised of those who were observed at follow-up ($R = 1$) remains representative for the population of interest and, if not, how estimates from the system comprising equations (2)–(6) can be used to correct for the sample selection bias caused by unit non-response.

Suppose we are interested in estimating in wave s ($s = 1, \dots, 6$) a conditional mean $E(Y_{is}|\boldsymbol{W}_{is})$ or the entire conditional distribution of Y. In the ELSA study, Y_{is} can be, for example a measure of disability or a SES outcome observed in wave $\boldsymbol{s}$. Since we can observe Y_{is} only for those who remain in the sample in wave $\boldsymbol{s}$, the simplest assumption we can make is that Y_{is} is *Missing Completely at Random* (MCAR), which implies independence between Y_{is} and R_{is}: $E(Y_{is}) = E(Y_{is}|\ R_{is} = 1)$. Under MCAR, we make the strong assumption that no other characteristics affect either Y_{is} or R_{is}.

A less restrictive assumption is *Missing at Random* (MAR) (see Little and Rubin (1987)) or *Selection on Observables* (Fitzgerald *et al.*, 1998). Under MAR, the missingness mechanism which affects $\boldsymbol{Y_{is}}$ does not depend on unobservables, conditional on observables ($\boldsymbol{W_{is}}$): $E(Y_{is}|\ \boldsymbol{W}_{is}) = E(Y_{is}|\ \boldsymbol{W}_{is}, R_{is} = 1)$. Decompose $\boldsymbol{W_{is}}$ into $(\boldsymbol{x_i}, h_i^*, e_i^*)$ and the MAR condition can be re-written as follows: $E(Y_{is}|\ \boldsymbol{W}_{is}) = E(\theta_{is}, R_{is}, Y_{is})$ with $\theta_{is} = \frac{1}{Pr(R_{is}=1|\boldsymbol{x}_{is}, h_i^*, e_i^*)}$ denoting the vector of individual weights which can be used to obtain consistent estimates of Y_{is}.[26] This procedure is the so-called "Inverse Probability Weights" (IPW) (Horvitz & Thompson, 1952).

It has been shown that, under certain conditions,[27] IPW is efficient in presence of exogenous covariates and it has been used extensively in recent years as the basis of many post-collection adjustment procedures aiming at attenuating, under MAR, sample selection bias.

For the core-member sample of ELSA, non-response weights were constructed by the ELSA team at household level using a propensity score weighting method. For wave 1, a logistic regression of retention in the study was used based only on age of the oldest household member, regional health authority, household size, social class, year of HSE interview and presence of long-standing illness as obtained in wave 0 (Taylor *et al.*, 2007).

By using auxiliary population information, a post-stratification adjustment is often made to ensure that the sample matches the population the study intends to represents. In the case of ELSA calibration weights were used for wave 1 to account for any potential bias caused by unequal selection probabilities (in the

---

[26] To see this, observe that: $E(\theta_{is}, R_{is}, Y_{is}) = E[E(\theta_{is}, R_{is}, Y_{is}|\boldsymbol{x}_{is}, h^*_{is}, e^*_{is}] = E[Pr(R_{is} = 1|\boldsymbol{x}_{is}, h^*_{is}, e^*_{is})\theta_{is}E(Y_{is}|\boldsymbol{x}_{is}, h^*_{is}, e^*_{is})] = E(Y_{is}|\boldsymbol{x}_{is}, h^*_{is}, e^*_{is}) = E(Y_{is})$ where $\boldsymbol{x}_{is}$ may contain time-invariant as well as time-variant characteristics. The latter can be observed only for those with $R_{is} = 1$. A usual assumption in longitudinal analysis is to instrument $\boldsymbol{x}_{is}$ with its lagged value (i.e. the values observed in wave $s-1$ of ELSA), available for both respondents and non-respondents of wave s.

[27] We refer to Wooldridge (2002) for a detailed presentation of IPW theory.

HSE) and refusals to be contacted for the ELSA study. Such further round of weights were constructed as the ratio of the sample size in the Census 2001 non-institutionalised population and ELSA, in cells defined by gender and age groups. Calibration weights were then multiply with the non-response IPW weights to better align the sample to the target population (Taylor *et al.*, 2007, pp. 45-46).

If non-response in wave $\boldsymbol{s}$ is totally random, using calibration weights that match the ELSA sample with census estimates would be enough to restore the representativity of ELSA at wave 1. If sample selection bias is determined only by factors included in the ELSA weights, then such weights are required for making inference at follows-up. But the issue of importance here is to see if sample selection is determined by additional factors that are not captured in the ELSA weights (but included in our application) that can lead to biased inference on the population of interest.

6.1 Implications of selection on observables for ELSA wave 1 tabulations

Table 6 shows the effect of weighting on cross-sectional distribution of a large number of SES, health and disability indicators, together with the proportion of those in receipt of disability-related benefits as observed in wave 1. We show four estimates. The first one refers to the unweighted statistics; the second corrects for non-response using ELSA wave 1 weights; and the third and fourth columns show sample statistics corrected by the inverse of the individual's fitted probability of retention from model D of Table 5 (*IPW-model D*, henceforth),[28] with and without applying calibration weights defined by the ELSA team. The third and fourth columns are meant to provide evidence of how far the alternative weights we propose affect statistics of interest.

The differences between unweighted and weighted sample statistics provides a measure of how selection on observables influences Y_i. In most instances, using weights has a considerable effect. Under the MAR assumption, the sample is older and in lower SES[29] than under MCAR. Moreover, weighted estimates pointed towards a slightly more disabled population with some important differences according to the weighting procedure in use.

By contrasting estimates in columns two and three of Table 6 we found samples which are similar in age but with important differences in term of SES and health status, being the sample weighted using *IPW-model D* less advantaged in term of SES and more functionally disabled – both in terms of reported Activity Daily Living (ADL) (Katz *et al.* (1963)) and Instrumental ADL (IADL) limitations, (Lawton and Brody (1969)) than the one weighted with ELSA original weights.

[28] It should be noticed that the standard IPW method does not involve latent variables. This analysis differs from previous studies allowing for the use of latent factors that have been constructed from estimates (i.e. not assumed directly observed).

[29] SES is measured by level of education, home-ownership, housing and financial wealth and in term of current income reported by the benefit units.

This sample also reflects a slightly higher prevalence of reported medical conditions, and ultimately, a higher percentage of individuals in receipt of the two main disability-related benefits: Attendance Allowance (AA) and Disability Living Allowance (DLA).

TABLE 6: Descriptive statistics ELSA wave 1 sample according different weighting procedures

indicator	Un-weighted sample (MCAR)		Weighted samples (MAR)					
			ELSA weights		IPW-model D		IPW-model D* calibration	
	mean	*sd*	mean	*sd*	mean	*sd*	mean	*sd*
Age of the respondent [a]	73.8	*6.702*	74.3	*6.927*	74.2	*6.856*	74.2	*6.843*
Married/partnered	0.590	*0.492*	0.580	*0.494*	0.580	*0.494*	0.575	*0.494*
Level of education								
No qualifications	0.387	*0.487*	0.388	*0.487*	0.382	*0.486*	0.385	*0.487*
lT high-school	0.312	*0.463*	0.324	*0.468*	0.334	*0.472*	0.333	*0.471*
High-school graduate	0.102	*0.303*	0.098	*0.297*	0.098	*0.297*	0.097	*0.296*
Some college	0.123	*0.329*	0.118	*0.323*	0.116	*0.320*	0.115	*0.319*
College and above	0.076	*0.265*	0.072	*0.259*	0.070	*0.255*	0.069	*0.254*
Home-ownership	0.763	*0.425*	0.754	*0.431*	0.753	*0.431*	0.753	*0.432*
Net housing wealth	148,617	*119,653*	147,974	*118,838*	147,461	*119,504*	147,469	*119,730*
Non-housing financial wealth	44,858	*91,515*	43,043	*88,920*	42,693	*89,294*	42,512	*88,733*
Total benefit unit income	13,708	*16,226*	13,427	*16,279*	13,310	*15,545*	13,275	*15,759*
Difficulties in ADLs								
none	0.716	*0.451*	0.714	*0.452*	0.704	*0.457*	0.704	*0.457*
1	0.138	*0.345*	0.141	*0.348*	0.139	*0.346*	0.138	*0.345*
2	0.065	*0.246*	0.066	*0.248*	0.067	*0.249*	0.067	*0.249*
3	0.035	*0.183*	0.036	*0.185*	0.036	*0.187*	0.037	*0.188*
4 or more	0.046	*0.210*	0.044	*0.206*	0.054	*0.227*	0.055	*0.227*
Difficulties in IADLs								
none	0.818	*0.386*	0.814	*0.389*	0.803	*0.398*	0.802	*0.398*
1	0.099	*0.299*	0.102	*0.302*	0.102	*0.303*	0.102	*0.303*
2	0.039	*0.194*	0.041	*0.199*	0.042	*0.200*	0.042	*0.201*
3	0.014	*0.117*	0.015	*0.123*	0.016	*0.126*	0.017	*0.128*
4 or more	0.030	*0.172*	0.028	*0.165*	0.037	*0.189*	0.037	*0.189*
Medical conditions								
high blood pressure	0.440	*0.496*	0.438	*0.496*	0.440	*0.496*	0.441	*0.497*
diabetes	0.094	*0.291*	0.092	*0.290*	0.097	*0.295*	0.096	*0.295*
cancer	0.077	*0.267*	0.078	*0.269*	0.074	*0.262*	0.075	*0.263*
lung disease	0.073	*0.259*	0.072	*0.258*	0.073	*0.261*	0.073	*0.259*
heart problems	0.253	*0.435*	0.252	*0.434*	0.255	*0.436*	0.253	*0.435*
stroke	0.069	*0.254*	0.070	*0.254*	0.073	*0.260*	0.071	*0.258*
arthritis	0.383	*0.486*	0.382	*0.486*	0.384	*0.486*	0.387	*0.487*
psychiatric problems	0.044	*0.205*	0.044	*0.204*	0.044	*0.206*	0.045	*0.206*
memory problems	0.008	*0.089*	0.008	*0.091*	0.009	*0.096*	0.009	*0.096*
Receipt of cash disability benefits (AA or DLA)	0.124	*0.330*	0.129	*0.336*	0.133	*0.340*	0.134	*0.340*

Notes: Author's computation based on the sample of respondents interviewed in both wave 0 *and* wave 1 (N=5,432). ([a]) collapsed at 90.

Sample statistics in column three are virtually identical to those in column four of Table 6. In other words, calibration weights modify only marginally the non-response weights, indicating limited bias due to the use of HSEs in defining the ELSA sampling frame.

Comparisons with administrative data are not straightforward, because they include the care home population.[30] Our weighting strategy gives an estimate of the number of AA/DLA recipients in England in 2002 about 80,000 higher than the one estimated using ELSA weights, and about 160,000 higher that the un-weighted estimate.[31] Our weighting strategy brings the estimate much closer to the figure in the official statistics.

Three main messages emerge from this analysis. First, as suggested in Taylor *et al.* (2007), descriptive analysis of the ELSA sample should be based on weighted data. Under the MAR assumption, the weighting strategy adopted does not play a very crucial role. However, our weighting approach tends to reduce SES statistics and increase prevalence of disability and receipt of related benefits, with potentially important consequences for evaluating reforms to the public system of care and support for disabled people. Finally, the small impact of calibration weights gives very little evidence of selection bias caused by the use of HSEs to draw the ELSA sample.

[30] We refer to Hancock *et al.* (2015) for details.

[31] Our figures refer to the non-institutionalised population in England, aged 65+.

6.2 Implications of selection on observables for ELSA wave 2 follow-up

Longitudinal studies such as the ELSA enable analysis of the dynamics of health. In what follows we extend previous analysis looking at the effects of sample selection at successive ELSA follow-up. This analysis requires re-estimating the system comprising equations (2)–(6), with R_i redefined as an indicator of retention (for eligible members only)[32] in *all* of the preceding waves up to the point of analysis. We do this up to wave 2 of ELSA only, but such analysis could be potentially extended up to the last wave currently available.

Structural parameters for the probability of remaining in the study in *both* waves 1 and 2 (conditional on being respondent in wave 0) are available in the Appendix Table A3.[33] The implications of the estimates are summarised in Appendix Figure A1 for the three illustrative 75-year-old individuals defined in section 4.

As before, fitted probabilities of retention are used to construct longitudinal weights for the subset of fully respondent members in wave 2. Tables 7 and 8

[32] We therefore excluded from the analysis 429 sample members who become ineligible for wave 2 of ELSA but we considered as refusals the 2,941 sample members who were not issued in wave 2 because all wave 1 respondents in the household explicitly refused to be recontacted. See sub-section 3.2 for details.

[33] They can be contrasted with structural parameters of retention for wave 1 only. With respect to estimates of model D in Table 5, we found a few important differences. The effect of the second spline of age is almost three times higher whereas living with a partner makes the probability of dropping out 1.4 times higher than the estimated effect of retention in wave 1. The signs of the educational attainments variables are in line with expectations and similar to those obtained in modelling retention for wave 1. Social class is now no longer associated with retention. While the coefficient associated with engagement is stable, the latent poor-health index increases considerably its significance in explaining retention, with its correlation with engagement virtually unaffected.

displays tabulations of some health indicators for the unweighted sample (column 1) and the weighted samples using ELSA weights available for wave 2 (column 2) and using our vector of weights built according estimates for Appendix Table A3.

Table 7 reports prevalence data for ELSA wave 2 respondents born on or before 1937. Weighted data indicate a higher prevalence of (I)ADL limitations and correspondingly higher percentage of individuals receiving disability benefits (DBs) than unweighted data, in particular when our set of weights is in use.

TABLE 7: Prevalence of (I)ADL limitations and receipt of DBs in wave 2 for those born in or before 1937

	No weights (MCAR)	*ELSA weights (MAR)*	IPW-model D*calibration weights
% by number of ADLs reported			
0	71.89	70.90	68.41
1	15.21	15.59	16.55
2	6.75	7.07	7.26
3	3.39	3.47	3.76
4	1.97	2.09	2.77
5+	0.80	0.88	1.24
% by number of IADLs reported			
0	81.97	80.62	77.92
1	9.99	10.45	11.16
2	4.53	4.86	5.44
3	1.27	1.45	1.60
4	1.12	1.32	1.50
5+	1.12	1.31	2.37
% in receipt of DBs			
No	87.34	86.49	84.85
Yes	12.66	13.51	15.15

Notes: Estimates based on the (weighted) sample of 3,955 individuals born in or before 1937 (aged 65+ in 2002), fully respondent (and without item non-response on relevant variables) in waves 0, 1 and 2 of ELSA as observed in wave 2.

Table 8 assesses the effects of different weighting strategies on estimates of the onset of new medical conditions[34] and functional limitations and the rate of new DBs receipt from wave 1 (2002/03) to wave 2 (2004/05). Overall, about 23.3% of the unweighted sample reported new medical conditions in wave 2 in addition to any reported in wave 1.

TABLE 8: Changes in reported medical conditions, functional limitations and receipt of DBs since last interview (wave 1) by age groups at baseline

	No weights (MCAR)	ELSA weights (MAR)	IPW-model D*calibration weights
Reported new medical conditions since wave 1			
65-74	22.2%	22.5%	22.3%
75-84	25.1%	25.4%	26.0%
85+	23.9%	24.0%	26.7%
overall	23.3%	23.6%	24.0%
Number of new medical conditions reported since wave 1			
65-74	1.16	1.16	1.16
75-84	1.20	1.20	1.20
85+	1.15	1.13	1.14
overall	1.17	1.17	1.18
Reported a change in functional limitations (ADLs) since wave 1			
65-74	12.6%	12.8%	13.4%
75-84	20.4%	20.6%	22.4%
85+	22.3%	20.8%	23.3%
overall	15.8%	16.1%	17.6%
Number of new ADLs reported since wave 1			
65-74	1.37	1.39	1.39
75-84	1.42	1.44	1.44
85+	1.64	1.52	1.54
overall	1.42	1.42	1.43
Reported being in receipt of DBs since last interview			
65-74	3.6%	3.8%	3.7%
75-84	9.3%	9.4%	9.9%
85+	17.3%	19.6%	23.1%
overall	6.2%	6.8%	7.6%

Notes: Estimates based on the (weighted) sample of 3,955 individuals born in or before 1937 (aged 65+ in 2002), fully respondent (and without item non-response on relevant variables) in waves 0, 1 and 2 of ELSA as observed in wave 2.

[34] The following medical conditions are considered: high blood pressure, diabetes, cancer, lung disease, heart problems, stroke, arthritis, psychiatric and memory problems.

Dividing the sample by age group,[35] we found that incidence of new medical conditions increase non-linearly both under MCAR and MAR, whereas they follow a more plausible increasing age-related trend using our IPW weights. All results show a significant selection on observables bias, which increases with age of the respondent.

We also found an increase of about 0.6% over a two-year period of the number of new DB recipients when MCAR assumption is relaxed in favour of MAR using ELSA weights. This difference increases to 1.4% when our weights were used. Again, the main contribution to this bias comes from the group of people aged 85+. Among them and over a two-year period, we estimated about 17.3% of new recipients under MCAR, +2.3% using ELSA weights and +5.8% using non-response weights we computed.

7 Discussion

This study investigated the existence of health-related non-response among eligible older members at the beginning of the English Longitudinal Study of Ageing (ELSA) and its consequences for assessing their health and receipt of disability–related benefits at follow-up.

[35] With age observed at baseline (year 2002).

Building on a rational choice approach, a latent factor structural equation approach was used – for the first time in our knowledge – to study the effect of individuals' health on the survey participation, controlling for socio-demographics and individuals' attitudes towards the survey.

The model uses a latent representation of individuals' health status, with self-reported indicators of health status assumed to be imperfect indicators of the underlying "true" health status. Individual "engagement" with the scope of the survey is also assumed to be a latent factor, measured by a range of proxy indicators and assumed to covary with individuals' observable characteristics and latent health.

We found evidence of a non-linear relationship between health and successive drop-out of eligible members from the ELSA study. In other words, while unhealthy individuals are less likely to remain in the study at follow-up, very healthy individuals do not show significantly lower dropout probabilities with respect to the remaining sample population. This is consistent with the view that while unhealthy members face higher physical and cognitive costs in participating, healthy individuals might have high concerns about the time cost of participating and in taking part in medical and cognitive tests which, apart from being time-consuming, may also be felt to be humiliating.

Despite the existence of a non-linear health-related retention bias, controlling for individuals' engagement increases the explanatory power of the survey participation model and dilutes the health gradient, the relevance of which, however, is not completely eliminated.

The "engagement" index we derived in fact played the largest and most significant role in explaining participation, providing empirical support for the idea that the collection of indicators of individuals' attitudes towards the survey (or interviewers' views concerning interviewees' attitudes) could lead to great gains at no excessive extra cost to the data collection agency. Such gains could be in the form of:

1. An early identification of sample members who are more at risk of dropping out at successive waves. We would perhaps be tempted to focus our efforts on maintaining participation of the subgroup of less engaged eligible members rather than on those in worse health status. The finding here, however, suggests that the less engaged are also those in worst health and therefore such a choice might be less competing than it would seem.

2. Designing specific instruments that could incentivise members' retention in the study. For instance, future research could perhaps considering ways of reducing costs of survey participation, particularly for those more at risk of dropping out (e.g. by conducting condensed, shorter interviews to obtain key information from those more at risk of dropping out) or by boosting the benefits of

participating (e.g. by finding effective ways of communicating medical feedback to the sample members from nurse visits, blood tests and other objective tests);

3. Better informing post-collection adjustment procedures that deal with the selective nature of non-participation, as we documented in the section 6 of this paper where we have investigated the extent to which initial sample selection in ELSA would introduce bias in univariate analyses aiming at describing the health status and disability benefit receipt at follow-up.

Our strategy was to investigate the performance of a new vector of non-response weights built upon our structural estimates to correct for potential selection bias under Missing at Random (MAR). Our study can be also seen as an assessment of whether dropouts at the beginning of ELSA are determined by additional factors not captured in the ELSA weights.

Under MAR, we documented a downward of the socio-economic status and an upward of disability estimates that would emerge by assuming Missing Completely at Random (MCAR). As a result, using our non-response weights we documented a larger (and then closer to official statistics) proportion of older people in receipt of disability-related benefits (Attendance Allowance and Disability Living Allowance) than the one obtained by assuming MCAR or by using ELSA weights for wave 1. More worryingly, we found that the selectivity of non-response behaviours is even more severe in wave 2 of ELSA, with important implications for making inference on the population of interest. On the other hand, as the use of calibration weights has shown, we found very little evidence of selection bias caused by the use of HSEs to draw the ELSA sample after conditioning non-response weights on a proper set of observable characteristics.

Additional research is needed. For example, it is required to better understand the determinants of such a significant loss of eligible sample members in wave 2 (see section 4.1) which seems mainly driven by the manifestation of a complete disinterest of *all* (eligible) household members to the study. Only then can the analysis be successfully extended up to the last wave currently available, making use of a fuller set of health and engagement indicators available in the ELSA follow-up. This would be of particular interest also for new longitudinal studies, such as the new Scottish Longitudinal Study of Ageing (THSLS).

As far as the econometric framework is concerned, this paper should be considered as a call for further research on this topic. Motivated by the desire to use – in a more structured way – as much as possible of the information available in survey data, it should not be advocated as a panacea but rather a possible alternative way to test theories and provide new insight into pre- and post- data collection processes.

References

Banks, J., Muriel, A., & Smith, J. P. (2010). *Attrition and Health in Ageing Studies: Evidence from ELSA and HRS*. IZA working paper (5161).

Beatty, P., & Herrmann, D. (2002). To Answer or Not to Answer: Decision Processes Related to Survey Item Nonresponse. In R. M. Groves, D. A. Dillman, J. L. Eltinge & R. J. Little (Eds.), *Survey Nonresponse* (pp. 71-85). New York: John Wiley and Sons.

Bhamra, S., Tinker, A., Mein, G., Ashcroft, R., & Askham, J. (2008). The Retention of Older People in Longitudinal Studies: A Review of the Literature. *Quality in Ageing and Older Adults, 9* (**4**), 27-35.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York: John Wiley and Sons.

Börsch-Supan, A., Hank, K., & Jürges, H. (2005). A New Comprehensive and International View on Ageing: Introducing the 'Survey of Health, Ageing and Retirement in Europe'. *European Journal of Ageing, 2* (**4**), 245-253.

Chatfield, M. D., Brayne, C. E., & Matthews, F. E. (2005). A Systematic Literature Review of Attrition between Waves in Longitudinal Studies in the Elderly Shows a Consistent Pattern of Dropout between Differing Studies. *Journal of Clinical Epidemiology, 58* (**1**), 13-19.

Colombo, F., Llena-Nozal, A., Mercier, J., & Tjadens, F. (2011). *OECD Health Policy Studies Help Wanted? Providing and Paying for Long-Term Care: Providing and Paying for Long-Term Care*: OECD Publishing.

Contoyannis, P., Jones, A. M., & Rice, N. (2004). The Dynamics of Health in the British Household Panel Survey. *Journal of Applied Econometrics, 19* (**4**), 473-503.

Copas, A. J., & Farewell, V. T. (1998). Dealing with Non-Ignorable Non-Response by Using an 'Enthusiasm-to-Respond'variable. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 161* (**3**), 385-396.

d'Uva, T. B., Lindeboom, M., O'Donnell, O., & Van Doorslaer, E. (2011). Slipping Anchor? Testing the Vignettes Approach to Identification and Correction of Reporting Heterogeneity. *Journal of Human Resources, 46* (**4**), 875-906.

Deaton, A. (2002). Policy Implications of the Gradient of Health and Wealth. *Health Affairs, 21* (**2**), 13-30.

Deaton, A. S., & Paxson, C. (2001). Mortality, Education, Income, and Inequality among American Cohorts. In D. A. Wise (Ed.), *Themes in the Economics of Aging* (pp. 129-165): University of Chicago Press.

Deeg, D. J. H., van Tilburg, T., Smit, J. H., & de Leeuw, E. D. (2002). Attrition in the Longitudinal Aging Study Amsterdam: The Effect of Differential Inclusion in Side Studies. *Journal of Clinical Epidemiology, 55* (**4**), 319-328.

DeMaio, T. J. (1980). Refusals: Who, Where and Why. *Public Opinion Quarterly, 44* (**2**), 223-233.

Dillman, D. A. (1978). *Mail and Telephone Surveys: The Total Design Method*. New York: Wiley Interscience.

Dunn, K. M., Jordan, K., Lacey, R. J., Shapley, M., & Jinks, C. (2004). Patterns of Consent in Epidemiologic Research: Evidence from over 25,000 Responders. *American Journal of Epidemiology, 159* (**11**), 1087-1094.

Fitzgerald, J., Gottschalk, P., & Moffitt, R. (1998). An Analysis of Sample Attrition in Panel Data - the Michigan Panel Study of Income Dynamic. *Journal of Human Resources, 33* (**2**), 251-299.

Graham, H. (2009). *Understanding Health Inequalities*: McGraw-Hill International.

Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy, 80* (**2**), 223-255.

Groves, R. M., Cialdini, R. B., & Couper, M. P. (1992). Understanding the Decision to Participate in a Survey. *Public Opinion Quarterly, 56* (**4**), 475-495.

Groves, R. M., & Couper, M. P. (1998). Nonresponse in Household Interview Studies. New York: John Wiley and Sons.

Hancock, R., Morciano, M., Pudney, S., & Zantomio, F. (2015). Do Household Surveys Give a Coherent View of Disability Benefit Targeting? A Multi-Survey Latent Variable Analysis for the Older Population in Great Britain. *Journal of the Royal Statistical Society. Series A, Statistics in Society, DOI: 10.1111/rssa.12107*.

Hawkes, D., & Plewis, I. (2006). Modelling Non-Response in the National Child Development Study. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 169* (**3**), 479-491.

Hill, D. H., & Willis, R. J. (2001). Reducing Panel Attrition: A Search for Effective Policy Instruments. *Journal of Human Resources*, 416-438.

Horvitz, D. G., & Thompson, D. J. (1952). A Generalization of Sampling without Replacement from a Finite Universe. *Journal of the American Statistical Association, 47* (**260**), 663-685.

Jenkins, S. P., Cappellari, L., Lynn, P., Jäckle, A., & Sala, E. (2006). Patterns of Consent: Evidence from a General Household Survey. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 169* (**4**), 701-722.

Johansson, F., & Klevmarken, A. (2006). Explaining the Size and Nature of Response in a Survey on Health Status and Economic Standard. *Journal of Official Statistics, 24* (**3**), 431-449.

Johnson, R. J., & Wolinsky, F. D. (1993). The Structure of Health Status among Older Adults: Disease, Disability, Functional Limitation, and Perceived Health. *Journal of Health and Social Behavior*, 105-121.

Jones, A. M., Rice, N., & Contoyannis, P. (2006). The Dynamics of Health. In A. M. Jones (Ed.), *The Elgar Companion to Health Economics* (pp. 17). Northampton: Edward Elgar Publishing, Inc.

Kapteyn, A., Michaud, P.-C., Smith, J., & van Soest, A. (2006). *Effects of Attrition and Non-Response in the Health and Retirement Study.* IZA working paper (2246).

Katz, S., Ford, A. B., Moskowitz, R. W., Jackson, B. A., & Jaffe, M. W. (1963). Studies of Illness in the Aged: The Index of Adl: A Standardized Measure of Biological and Psychosocial Function. *JAMA, 185* (**12**), 914-919.

Kho, M. E., Duffett, M., Willison, D. J., Cook, D. J., & Brouwers, M. C. (2009). Written Informed Consent and Selection Bias in Observational Studies Using Medical Records: Systematic Review. *BMJ, 338*.

Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling*: Guilford press.

Laurie, H., Smith, R., & Scott, L. (1999). Strategies for Reducing Nonresponse in a Longitudinal Panel Survey. *Journal of Official Statistics, 15* (**2**), 269-282.

Lawton, M., & Brody, E. (1969). Physical Self-Maintenance Scale (Functional Assessment). *Gerontologist, 9*, 179-186.

Lee, L.-F. (1982). Health and Wage: A Simultaneous Equation Model with Multiple Discrete Indicators. *International Economic Review, 23* (**1**), 199-221.

Lepkowski, J. M., & Couper, M. P. (2002). Nonresponse in the Second Wave of Longitudinal Household Surveys. In R. M. Groves, D. A. Dillman, J. L. Eltinge & R. J. Little (Eds.), *Survey Nonresponse* (pp. 259-272). New York: John Wiley and Sons.

Little, R. J., & Rubin, D. B. (1987). *Statistical Analysis with Missing Data*. New York: John Wiley and Sons.

Liu, K. (1988). Measurement Error and Its Impact on Partial Correlation and Multiple Linear Regression Analyses. *American Journal of Epidemiology, 127* (**4**), 864-874.

Loosveldt, G., Pickery, J., & Billiet, J. (2002). Item Nonresponse as a Predictor of Unit Nonresponse in a Panel Survey. *Journal of Official Statistics, 18* (**4**), 545-558.

Lynn, P., & Clarke, P. (2002). Separating Refusal Bias and Non-Contact Bias: Evidence from Uk National Surveys. *Journal of the Royal Statistical Society. Series D (The Statistician), 51* (**3**), 319-333.

Matthews, F., Chatfield, M., Brayne, C., & Medical Research Council Cognitive Function Ageing Study. (2006). An Investigation of Whether Factors Associated with Short-Term Attrition Change or Persist over Ten Years: Data from the Medical Research Council Cognitive Function and Ageing Study (MRC CFAS). *BMC Public Health, 6* (**1**), 185.

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus: User's Guide. Seventh Edition*. Los Angeles, CA: Muthén & Muthén

Nicoletti, C., & Buck, N. (2004). *Explaining Interviewee Contact and Co-Operation in the British and German Household Panels.* ISER Working Paper Series (6).

Nicoletti, C., & Peracchi, F. (2005). Survey Response and Survey Characteristics: Microlevel Evidence from the European Community Household Panel. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 168* (**4**), 763-781.

Norris, F. H. (1985). Characteristics of Older Nonrespondents over Five Waves of a Panel Study. *Journal of Gerontology, 40* (**5**), 627-636.

Pyy-Martikainen, M., & Rendtel, U. (2008). Assessing the Impact of Initial Nonresponse and Attrition in the Analysis of Unemployment Duration with Panel Surveys. *AStA Advances in Statistical Analysis, 92* (**3**), 297-318.

Scholes, S., Taylor, R., Cheshire, H., Cox, K., & Lessof, C. (2008). *Retirement, Health and Relationships of the Older Population in England: The 2004 English Longitudinal Study of Ageing Technical Report*. London: National Centre for Social Research (Natcen).

Stoop, I. A. L. (2005). *The Hunt for the Last Respondent: Nonresponse in Sample Surveys*. The Haque: SCP, Social and Cultural Planning Office of the Netherlands.

Taylor, R., Conway, L., Calderwood, L., Lessof, C., Cheshire, H., Cox, K., & Scholes, S. (2007). *Health, Wealth and Lifestyles of the Older Population in England: The 2002 English Longitudinal Study of Ageing Technical Report*. London: National Centre for Social Research (Natcen).

Tinker, A., Mein, G., Bhamra, S., Ashcroft, R., & Seale, C. (2009). Retaining Older People in Longitudinal Research Studies: Some Ethical Issues. *Research Ethics Review, 5* (**2**), 71-74.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The Psychology of Survey Response*: Cambridge University Press.

Uhrig, S. (2008). *The Nature and Causes of Attrition in the British Household Panel Study.* ISER Working Paper Series (5).

Van Beijsterveldt, C. E. M., van Boxtel, M. P. J., Bosma, H., Houx, P. J., Buntinx, F., & Jolles, J. (2002). Predictors of Attrition in a Longitudinal Cognitive Aging Study: The Maastricht Aging Study (MAAS). *Journal of Clinical Epidemiology, 55* (**3**), 216-223.

Van Vliet, R. J., & Van Praag, B. (1987). Health Status Estimation on the Basis of Mimic-Health Care Models. *Journal of Health Economics, 6* (**1**), 27-42.

Verbrugge, L. M., & Jette, A. M. (1994). The Disablement Process. *Social Science and Medicine, 38* (**1**), 1-14.

Wallace, R. B., & Herzog, A. R. (1995). Overview of the Health Measures in the Health and Retirement Study. *Journal of Human Resources, 30*, S84-S107.

Wang, J., & Wang, X. (2012). *Structural Equation Modeling: Applications Using Mplus*: John Wiley & Sons.

Watson, D. (2003). Sample Attrition between Waves 1 and 5 in the European Community Household Panel. *European Sociological Review, 19* (**4**), 361-378.

Watson, N., & Wooden, M. (2009). Identifying Factors Affecting Longitudinal Survey Response. In P. Lynn (Ed.), *Methodology of Longitudinal Surveys* (pp. 157-181). Chichester, UK: John Wiley & Sons, Ltd.

Wolfe, B. L., & Behrman, J. R. (1984). Determinants of Women's Health Status and Health-Care Utilization in a Developing Country: A Latent Variable Approach. *The Review of Economics and Statistics, 66* (**4**), 696-703.

Wooldridge, J. M. (2002). Inverse Probability Weighted M-Estimators for Sample Selection, Attrition, and Stratification. *Portuguese Economic Journal, 1* (**2**), 117-139.

Zabel, J. E. (1998). An Analysis of Attrition in the Panel Study of Income Dynamics and the Survey of Income and Program Participation with an Application to a Model of Labor Market Behavior. *Journal of Human Resources*, 479-506.

Appendix

A generalised SEM of survey participation with q-latent variables

We allow the integration of latent factors in the general framework used in modelling survey participation. Let S_i^* denote the net utility an individual derives from selecting a certain decision in regards participation to the survey and R_i denote a binary variable indicating his or her actual decision (so $R_i = 1$ if the individual remains in the study and $R_i = 0$ otherwise). Assuming utility maximisation:

$$R_i = 1 \; if \; S_i^* \geq 0, R_i = 0 \; otherwise \tag{A1}$$

Following notation in Bollen (1989), the net utility S_i^* can be modelled as:

$$S_i^* = \boldsymbol{\alpha} \boldsymbol{z_i} + f(\boldsymbol{\xi_i}; \boldsymbol{\gamma}) + \varepsilon_i \tag{A2}$$

where $\boldsymbol{z_i}$ denote a vector of exogenous predictors of survey participation and $\boldsymbol{\alpha}$ the associated coefficients; $\boldsymbol{\xi_i}$ is a vector of q-latent exogenous variables with manifested indicators denoted by $\boldsymbol{w}$ and $\boldsymbol{\gamma}$ is a vector of fixed coefficients reflecting the importance of each latent factor on S_i^*. The mapping of manifested indicators to the q-latent factors is accomplished by a general measurement equation:

$$w_k = \mathbf{1}\Big(\lambda_{kq} \boldsymbol{\xi_q} + \delta_k > 0\Big) \tag{A3}$$

where w_k is a random vector of manifested indicators $(k = 1,2, \dots K)$, λ_{kq} are factor loadings that map manifested indicator k on to latent factor q through the indicator function $\mathbf{1}(.)$, and $\boldsymbol{\xi}_q$ and δ_k are latent factors and errors, respectively, which are assumed to be uncorrelated. Note that this model can be easily extended to include a mixture of continuous and ordered indicators in w. For identification purposes, given that the scales of the q latent variables are arbitrary, we need normalising them by setting either a factor loading λ_{1q} to be equal to 1 or the residual variance of $\boldsymbol{\xi}_q$ to be equal to 1.

Because only $\boldsymbol{w}$ and R can be observed, any inference must be based on the joint distribution whose density can be written generally as:

$$P(R_i = 1, \boldsymbol{w}|\boldsymbol{z}; \theta) = \int \quad P(R_i = 1|\boldsymbol{z}; \boldsymbol{\xi}; \boldsymbol{\alpha}, \boldsymbol{\gamma})\, g(w|\boldsymbol{\xi}; \boldsymbol{\lambda}, \boldsymbol{\Psi}) h(\boldsymbol{\xi}) d\boldsymbol{\xi} \qquad \text{(A4)}$$

where $\int$ is the support of $\boldsymbol{\xi}$. The first term of the integrand is the retention probability conditional on the latent variables $\boldsymbol{\xi}$, where $\boldsymbol{z}$ and $\boldsymbol{\alpha}$ denote again the (error-free) explanatory variables in the retention equation that may influence survey participation and the associated coefficients, respectively; the second term $g(w|\boldsymbol{\xi}; \boldsymbol{\lambda}, \boldsymbol{\Psi})$ corresponds to the measurement equation and is the conditional distribution of the manifested items w given the latent variables $\boldsymbol{\xi}$.

The vector θ represents the model parameters $(\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\lambda}, \boldsymbol{\Psi})$, with $\boldsymbol{\Psi}$ being the covariance of the random disturbance terms of Equation A3.

Notice that Equation A4 assumes that, conditional on the latent variables $\boldsymbol{\xi}$, the retention probability $P(R_i = 1|\boldsymbol{z}; \boldsymbol{\xi}; \boldsymbol{\alpha}, \boldsymbol{\gamma})$ and the distribution of the indicators $g(w|\boldsymbol{\xi}; \boldsymbol{\lambda}, \boldsymbol{\Psi})$ are independent; that is, the joint distribution of the two can be given by the product of the marginals.

Model parameters θ can be estimated by means of maximum likelihood procedure, maximising the following:

$$L(.) = \sum_i P(R_i = 1, \boldsymbol{w}|\boldsymbol{z};\theta) =$$

$$\sum_i \int_N P(R_i = 1|\boldsymbol{z};\boldsymbol{\xi};\boldsymbol{\alpha},\boldsymbol{\gamma})\, g(w|\boldsymbol{\xi};\boldsymbol{\lambda},\boldsymbol{\Psi}) h(\boldsymbol{\xi}) d\boldsymbol{\xi} \tag{A5}$$

In developing the basic formulation we have described the survey participation model as being influenced by exogenous latent variables $\boldsymbol{\xi}$. It should be noticed that Equation A3 contains only the latent variables on the right-hand-side. However, they may also contain individual characteristics which might capture systematic response bias when the individual is providing response to the indicators. Moreover, note that $f(.)$ in Equation A2 is deliberately undefined. Typically, the function is specified to be linear in its parameter, but this is not necessary. Finally, note that the distribution of the error terms must be specified, leading to additional unknown covariance parameters.

Robustness checks: uncorrelated latent factors

Table A1: Structural parameters for the probability of retaining the study in wave 1: Model D no correlation between h^* and e^*

Covariate	Latent poor-health equation		Latent engagement equation		Retention equation	
	Coefficient	*S.E*	*Coefficient*	*S.E*	*Coefficient*	*S.E*
Female	0.050**	0.023	-0.184***	0.032	-0.001	0.032
Spline age 65-74	-0.006***	0.001	-0.011	0.008	0.01	0.008
Spline from age 74+ [a]	0.017***	0.003	-0.030***	0.005	-0.012**	0.006
Married/cohabiting	0.032	0.028	0.106**	0.047	-0.177***	0.048
Completed education before 14 years old	0.185***	0.028	-0.038	0.047	-0.094**	0.048
Completed education after 19 years old	-0.244***	0.051	-0.122	0.082	0.206**	0.082
Non-white	0.542***	0.083	-0.611***	0.109	0.167	0.136
Home-owner	-0.301***	0.031	0.095*	0.051	-0.069	0.052
Social class: manual worker	0.177***	0.027	-0.131***	0.043	-0.037	0.044
h^*					-0.046**	0.02
$(h^*)^2$					-0.025*	0.015
e^*					0.734***	0.042

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Estimates obtained from the sample composed by 5,050 core member individuals (65+) interviewed in wave 0 and eligible for wave 1 and wave 2. (a) collapsed at 90. All models also include dummy variables on region of residence, if living in urban area (>10K), and from which HSE cross-section the eligible person was drawn. Standard errors were clustered at household level.

Robustness checks: SEM vs. single-equation approach

Estimates in Table 5 can be contrasted with the reduced-form estimates commonly used in modelling survey participation. These, available in the Appendix Table A2, are obtained by fitting maximum-likelihood probit models. The first model includes controls for socio-demographic characteristics only but we have progressively included $\boldsymbol{H}$ (model 2) and $\boldsymbol{E}$ (model 3), treated as exogenous (error-free) explanatory variables.

A reduced-form approach is correct and not without interest if we are concerned with the "ultimate" determinants of survey participation rather than isolating the effect of $\boldsymbol{x}$ on health and engagement. This approach yields similar results the one displayed in Table 5 in the sense that corresponding coefficient estimates have the same sign. However, the standard errors are generally lower the one obtained using a SEM approach, yielding higher significance level for some covariates. This is because a reduced form aims at capturing the "total" effect that $\boldsymbol{x}$ would have on survey participation, whereas a "structural" framework allows to separate the direct impact that $\boldsymbol{x}$ has in explain survey participation from the indirect impact that $\boldsymbol{x}$ would have via its gradient with health and engagement.

Female gender, age and low SES are associated with lower retention in wave 1 for all reduced-form specifications. The structural approach, however, revealed that gender and social class play only an indirect role in explaining survey retention, with most of their effects mediated by the latent variables. Only three of the fourteen health indicators play a significant role (model 2), due to the high multicollinearity among $\boldsymbol{H}$ indicators. The inclusion of $\boldsymbol{E}$ as covariate (model 3) reduces further the significance of $\boldsymbol{H}$ indicators. However, in this reduced-form approach, the most influential predictors of non-response are those which were most influential in determining e^* in the structural equation approach.

Table A2: Parameter estimates for proxy-variables models used to predict individual's retention in the study in wave 1

Covariate	Model 1	Model 2	Model 3
Female	-0.109***	-0.114***	-0.081***
Spline age 65-74	0.002	0.001	0.004
Spline from age 74+ [a]	-0.027***	-0.027***	-0.023***
Married/cohabiting	-0.080*	-0.077*	-0.112**
Completed education before being 14 years old	-0.104**	-0.101**	-0.096**
Completed education after being 19 years old	0.092	0.081	0.130*
Non-white	-0.252*	-0.218*	-0.093
Home-owner	0.012	0	-0.025
Social class: manual worker	-0.111***	-0.097**	-0.088**
Self-reported health status (1 very good;...; 5 very bad)		-0.087***	-0.064***
Infectious disease		-0.015	-0.044
Neoplasms & benign growths [a]		0.087	0.053
Endocrine & metabolic		-0.065	-0.066
Blood & related organs		0.026	0.052
Mental disorders		0.05	-0.005
Nervous system		0.107	0.084
Eye complaints		0.056	0.029
Ear complaints		0.062	0.029
Heart & circulatory system		0.177***	0.161**
Respiratory system		0.011	-0.06
Digestive system		-0.069	-0.096
Genito-urinary system		0.077*	0.06
Musculoskeletal system		0.035	0.048
Consent link survey data with admin data			0.663***
Complete/return the self-completion booklet			0.257**
Respondent to financial questions (income sources and savings)			0.289**
Consent having a nurse visit			0.826***
constant	0.674	0.823	-1.256*
N		8,420	
Log-likelihood	-5389.101	-5367.72	-5018.958
Pseudo-R2	0.018	0.022	0.086
AIC	10820.203	10805.44	10115.916
BIC	10968.009	11051.783	10390.412

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. ([a]) collapsed at 90. Estimates in the table refers to 3 different probit specifications (see section 4) of the determinants of retaining the ELSA study in wave 1. The model also include dummy variables on region of residence, if living in urban area (>10K), and from which HSE cross-section the eligible person was drawn. Standard errors were clustered at household level.

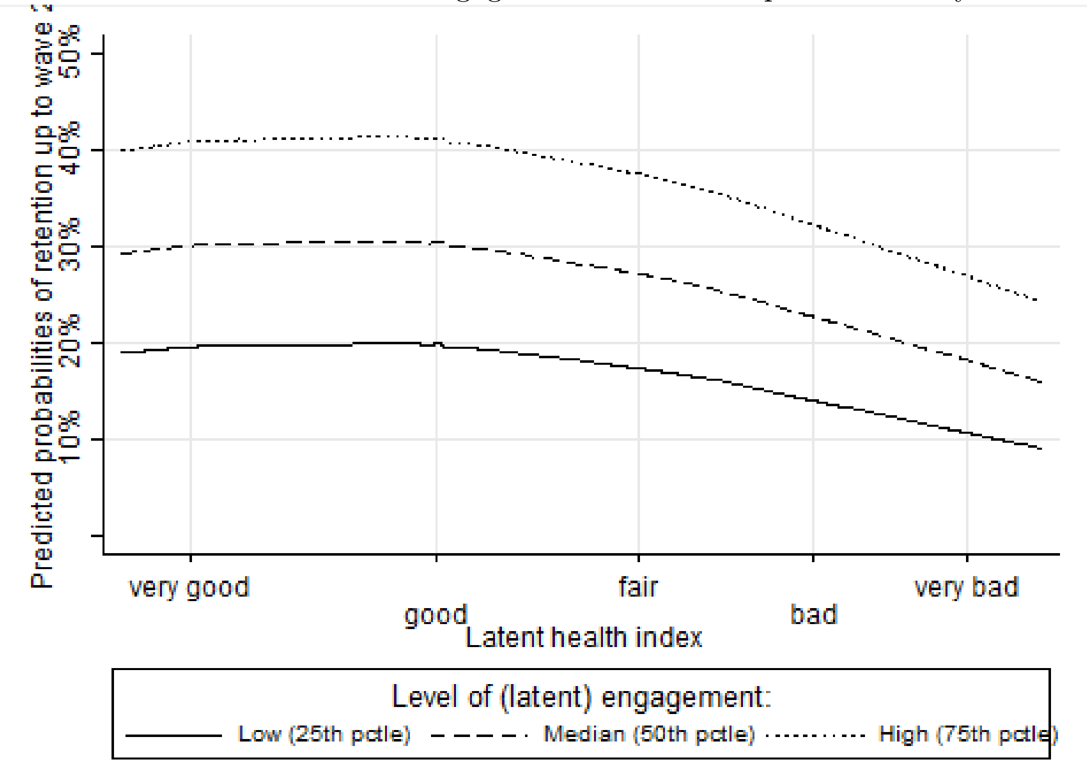
Modelling survey-participation in both waves 1 and 2

Table A3: Structural parameters for the probability of retaining the study in both waves 1 and 2

Covariate	Latent poor-health equation		Latent engagement equation		Retention equation	
	coefficient	S.E	coefficient	S.E	coefficient	S.E
Female	0.059**	0.024	-0.242***	0.047	0.019	0.033
Spline age 65-74	-0.007***	0.001	-0.017	0.012	0.012	0.009
Spline from age 74+ [a]	0.017***	0.003	-0.048***	0.008	-0.031***	0.006
Married/cohabiting	0.027	0.029	0.154**	0.067	-0.250***	0.049
Completed education before 14 years old	0.185***	0.028	-0.038	0.067	-0.083*	0.048
Completed education after 19 years old	-0.244***	0.052	-0.167	0.119	0.249***	0.085
Non White	0.572***	0.085	-0.888***	0.156	-0.015	0.137
Home owner	-0.291***	0.033	0.180**	0.074	0.028	0.054
Social class: manual worker	0.178***	0.027	-0.184***	0.062	-0.135***	0.045
h^*					-0.088***	0.023
$(h^*)^2$					-0.046***	0.016
e^*					0.511***	0.031
$cov(h^*, e^*)$					-0.073**	0.031
Free parameters			104			
Log-likelihood			-46883.195			
Correction for non-normality factor			1.1234			
AIC			93974.39			
BIC			94370.45			

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Estimates obtained from the sample composed by 7,791 individuals born in or before 1937 (aged 65+ in 2002), fully interviewed in wave 0 and eligible for wave 1 and wave 2 with no missing values at the indicators used for the analysis. (a) collapsed at 90. All models also include dummy variables on region of residence, if living in urban area (>10K), and from which HSE cross-section the eligible person was drawn. Standard errors were clustered at household level.

Figure A1: Predictions of the probability of remaining in the ELSA study up to wave 2 by latent health index for three 75-old representative individuals with different level of engagement with the scope of the study



Notes: Predicted probabilities computed using estimates in Appendix Table A3. To facilitate the interpretation, the labels for the x-axis in the figure refer to the categories of the 5-scale SAH indicator (the most influential factor in determining latent health) that corresponds to the mean values of the Empirical Bayes prediction of latent health observed at the 5 categories of the SAH question.

Chapter 2:

Do household surveys give a coherent view of disability benefit targeting? A multi-survey latent variable analysis for the older population in Great Britain*

Abstract: We compare three major UK surveys, BHPS, FRS and ELSA, in terms of the picture they give of the relationship between disability and receipt of the Attendance Allowance (AA) benefit. Using the different disability indicators available in each survey, we use a structural equation approach involving a latent concept of disability in which probabilities of receiving AA depend on disability. Despite major differences in design, once sample composition is standardised through statistical matching, the surveys deliver similar results for the model of disability and AA receipt. Provided surveys offer a sufficiently wide range of disability indicators, the detail of disability measurement appears relatively unimportant.

Keywords: disability indices, disability benefits, multiple surveys.

JEL codes: C81, I18, I38.

---

* This chapter has been published in the *Journal of Royal Statistical Society, series A: Statistics in Society* and it is a joint work with Ruth Hancock, Stephen Pudney and Francesca Zantomio. An earlier version is available as HEG working paper 13-03: https://www.uea.ac.uk/medicine/health-economics-group/working-papers. S. Pudney conceptualized ideas and supervised all aspects of its implementation. M. Morciano derived the dataset from the English Longitudinal Study of Ageing (ELSA), conducted the statistical analysis and synthesized and interpreted findings. R. Hancock derived the dataset from the Family Resources Survey (FRS) whereas F. Zantomio derived the dataset from the British Household Panel Survey (BHPS). All authors contributed to the writing of the article and reviewing drafts.

Data from the FRS are made available by the UK Department of Work and Pensions through the UK Data Archive. Material from the FRS is Crown Copyright and is used by permission. The BHPS data used were originally collected by MiSoC at the University of Essex (now incorporated within the Institute for Social and Economic Research) and are made available through the UK Data Archive (UKDA). Data from the ELSA were developed by researchers based at University College London, the Institute for Fiscal Studies and the National Centre for Social Research (NatCen) and are made available through the UKDA. NatCen provided geographical indicators. Neither the collectors of the data nor the UKDA bears any responsibility for the analyses or interpretations presented here.

1 Introduction

Developed countries like the UK will face severe problems in supporting the projected future growth in the disabled population (McVicar 2008), and in the older disabled population in particular (Karlsson *et al.* 2006, OECD 2005, Pickard *et al.* 2007). In the UK, there has been a long series of policy reviews by a Royal Commission (Sutherland 1999), the independent King's Fund (Wanless 2006), the government (Department of Health 2009), the Commission on Funding of Care and Support (CFCS 2011) and various parliamentary select committees. The current UK government has recently announced changes to some aspects of the long-term care funding system (Department of Health 2013) but debate continues on how best to provide public support to older people with care needs. Such debate and associated policy reform should ideally be evidence-based. This requires a robust and accurate baseline picture of the distribution of support for people with disabilities, allowing the development of statistical models to project changes in this picture as disability levels rise and alternative policy structures are implemented. In turn, this requires good survey data on patterns of disability and receipt of support.

The importance of disability as a policy issue is matched only by the vast range of survey questions that have been used to measure it, and the proliferation of disability indicators across surveys presents difficulties for empirical research.

There are many available question designs, supported by limited testing of external validity, internal consistency and test-retest reliability, and some cognitive evaluation of specific question designs (see Sturgis *et al.* 2001 and Jagger *et al.* 2009 for reviews of UK surveys). It is widely recognised that any particular set of disability indicators may give an imperfect description of the concept of disability relevant to the analysis and that bias may result from neglect of the measurement error problem (Bound 1991). However, there has been little cross-survey comparative work which considers the consistency of the empirical 'story' that policy-makers would get from surveys offering different sets of disability indicators. In practice, researchers often use disability indicators that happen to be available in a survey chosen for convenience or to meet other requirements, and the robustness issue is rarely considered systematically. The Green Paper (Department of Health 2009), State of the Nation' report (Cabinet Office 2010) and the report of the Commission on Funding Care and Support (CFCS 2011) are examples of policy documents based on research using a mixture of different survey sources for different purposes.

For policy purposes, we are interested not only in the measurement of disability, but also in its relationship with other key variables like receipt of public support. In this study, we focus on a particular form of public support: the disability-linked cash benefits which are available to older people. The main disability benefit for people aged 65 or over in the UK is Attendance Allowance (AA), which is administered by the Department for Work and Pensions (DWP) and designed to help meet the extra costs arising from disability. Besides the age restriction, eligibility for AA requires the claimant to be in need of care in order to perform daily activities. The AA claim form says “you may get Attendance Allowance if your disability means that you need help with your personal care or you need someone to supervise you for your own or someone else’s safety”. It defines help with personal care as “day-to-day help with things like washing (or getting in or out of the bath or shower), dressing, eating, going to and using the toilet, or telling people what you need or making yourself understood”; and supervision as needing “someone to watch over you to help you avoid substantial danger to yourself or other people” (Department for Work and Pensions, 2013). The benefit is not means tested and (in 2012/13) is worth either £51.85 per week, if care is needed during either day or night, or £77.45, if care is needed during both. Eligibility for AA is difficult to assess from survey data. In practice, decisions on claims are made by programme administrators on the basis of claimants’ reported health problems and consequent care needs. Once the claim is made, written evidence is examined by administrative assessors, who can require a medical examination of the claimant. An element of judgement is inevitable, so eligibility is uncertain, even with access to the same information as the administrative assessor. A further challenge is that the information on which the award decision is made is not observable directly in survey data. Rather, surveys offer a set of disability-related eligibility indicators, from which inference on the success of disability targeting must be drawn. AA is assessed solely on an individual's need for care. It is not means-tested (nor taxable), and is unaffected by the presence or circumstances of other household members. So it is possible for more than one household member to receive AA.

Our policy motivation has implications for the appropriate conceptualisation of disability. We are not concerned here with medical concepts of impairment, but rather disability conceived as a set of constraints on functioning which originate from health impairments broadly defined. This corresponds to Sen's (1985) "capabilities" approach, which sees the individual choosing a consumption vector x from a choice set X and a pattern of commodity utilisation $f(.)$ from a set of possible utilisation functions F. The individual's chosen vector of "functionings" is $b = f(x)$, which is thus constrained by his or her economic entitlements (X) and available ways of using economic resources (F). We view the concept of disability as a health-related limitation on the set F relative to some socially-agreed minimal norm N. The aim of disability policy is to offer support to people for whom $F \subset N$. Support may take the form of cash or services, both of which expand the individual's choice set X, and it may be universal, in which case support is independent of the pre-intervention X, or means-tested, in which case entitlement depends on X. The important point here is that the concept of disability is concerned with constraints on basic functionings, rather than medical conditions themselves. The survey indicators used to measure disability should therefore focus on potential difficulties with everyday activities rather than health or disease.

The contribution of this paper is to investigate whether different indicators of disability, collected in three widely-used household surveys, are consistent with a common set of findings relating to the targeting of disability benefits for older people. If we admit the possibility that underlying disability is multi-dimensional, there are two aspects to this comparability issue: completeness and compatibility. A survey is *complete* in its coverage of disability if its questionnaire content generates disability indicators that are capable of reflecting all the multiple dimensions of disability. Two surveys are mutually *compatible* if their respective indicators of any particular dimension of disability give the same undistorted picture of that underlying concept. For researchers using similar methods but different data sources to be sure of agreeing on their conclusions, both completeness and compatibility are necessary in general. We investigate three British surveys, the Family Resources Survey (FRS), the English Longitudinal Survey of Ageing (ELSA) and the British Household Panel Survey (BHPS), which have been widely used for research on health, disability and related topics. We find that compatibility is not a serious difficulty, although there are some signs that completeness is a problem for the BHPS.

Typically, the statistical analysis of disability benefit receipt employs a single-equation framework, in which a variety of disability indicators (or a count index of them) are used as explanatory covariates, together with several other characteristics related to socio-economic status (SES) (see Berthoud and Hancock 2008, Forder and Fernandez 2009 and Zantomio 2013 for examples). Instead, we use a structural equation approach involving a latent concept of disability to study the relationships between disability status, SES characteristics, and receipt of AA in the BHPS, ELSA and the FRS, at (almost) a single time point, 2002/03. We assume that an individual's disability status is not directly observable but reflected in varying degrees by members of a set of imperfect but observable survey indicators. In this respect we follow a number of authors since Lee (1982), Van de Ven and Van der Gaag (1982) and Wolfe and Behrman (1984) in considering health status as a latent concept. We assume that the underlying latent disability measure (η) is influenced by a set of SES characteristics and the probability of receiving AA is a function of η and SES characteristics. See Bollen (1989) for a review of this class of latent variable simultaneous equation models.

This methodological approach has two major advantages. First, overcoming the arbitrariness of approaches based on a limited set of disability indicators or a scalar (usually unweighted) count of them, the latent variable framework allows us to develop an index of disability which makes use of all available sample information. This composite index can then be used as a sounder basis for policy analysis focused on the targeting of disability benefit. Second, the latent variable framework reduces the scope for bias arising from the measurement error in observed disability-related indicators and therefore gives more reliable estimates of the relationship between benefit receipt and influences like disability and income – again improving the robustness of an analysis of benefit targeting. To our knowledge the latent disability approach has not been applied in multiple surveys each with different indicators of disability and the application to disability benefit receipt is also novel.

In sections 2 and 3 of the paper, we describe the methodological framework and the three surveys, documenting the distributional characteristics of the variables used. Results from the model fitted to the full (unmatched) samples are discussed in Section 4. Statistical models are best seen as local approximations, so comparison of evidence from different surveys may be influenced by differences in sample composition as well as the design of survey instruments. In section 5 we discuss ways of harmonising the samples, and opt for matching techniques to obtain samples with a (near-) common distribution for the SES covariates. This reduces the scope of the comparison slightly (the common support constraint) but has the advantage of removing differences due to model approximation errors at the periphery of the region covered by the survey samples. In section 6, we establish the robustness of our findings by examining their sensitivity to various aspects of the analytical approach.

2 A latent structural model of disability status and benefit receipt

In the gerontology literature, Johnson and Wolinsky (1993) conceptualise the dynamics of health status in the older population, viewing functional limitations as outcomes of latent disability. Consistent with this view, we model 'true' disability status as an unobservable, possibly multidimensional, phenomenon, which is influenced by socio-economic characteristics and circumstances. We observe a set of survey indicators, each of which provides a 'noisy' measure of underlying disability, satisfying the classical measurement error assumption that all correlation with other socio-economic characteristics is explained by latent disability. The main outcome of interest, receipt of AA, depends on latent disability and the set of socioeconomic characteristics which influence an individual's propensity to claim and be awarded AA.

Analysis is based on independent samples of n^s individuals in surveys $s = 1, 2, 3$. Each sampled individual i is characterised by: unobserved Q-dimensional 'true' disability $\boldsymbol{\eta}_i = (\eta_{i1} \dots \eta_{iQ})$; socio-economic individual characteristics $\boldsymbol{Z}_i$ observable in all surveys; a set of survey-specific disability-related discrete indicators D^s_{ij} , $j = 1 \dots J^s$; and a binary indicator of benefit receipt ($R_i = 1$) or non-receipt ($R_i = 0$). We aim to draw inferences about the conditional distributions $P(\boldsymbol{\eta}|\boldsymbol{Z})$ and $P(R|\boldsymbol{\eta}, \boldsymbol{Z})$ which describe respectively the distribution of disability in the population and the relationship between benefit receipt and the individual's disability and other characteristics. By definition, these population distributions are independent of any survey used to draw inferences about them. An important question is whether the distributions $P^s(R, D_1^s \dots D_{J^s}^s | \boldsymbol{Z})$ produced by the three surveys with their different disability indicators nevertheless give a coherent indication of underlying 'true' disability $\boldsymbol{\eta}$ and its relationship with benefit receipt R.

We estimate a Structural Equation Model (SEM) which comprises three components: a measurement model, a disability model and a benefit receipt model. We use an ordinal quasi-linear structure for disability measurement:

$$\widetilde{D}_{ij}^s = \alpha_j^s + \lambda_{j1}^s \eta_{i1} + \cdots + \lambda_{jQ}^s \eta_{iQ} + \varepsilon_{ij}^s \tag{1}$$

$$D_{ij}^s = m \qquad \text{iff} \qquad A_{jm-1}^s \le \widetilde{D}_{ij}^s < A_{jm}^s , \qquad m = 1, \dots, M_j^s \tag{2}$$

where: the coefficients λ_{jq}^s are factor loadings relating observed indicators in survey s to underlying disability; ε_{ij}^s is a normally-distributed residual term representing random response error, implying an ordered probit link function generating the observable indicator D_{ij}^s from its unobserved continuous form $\widetilde{D}_{ij}^s$. M_j^s is the number of response categories for indicator D_{ij}^s and the A_{jm}^s are threshold parameters. In the following we refer to equations (1) and (2) as the measurement model. The qth disability component η_{iq} is related to $\boldsymbol{Z}_i$ through a linear relationship representing the processes leading to disability (disability model):

$$\eta_{iq} = \boldsymbol{\theta}_q \boldsymbol{Z}_i + \upsilon_{iq} \tag{3}$$

where $\boldsymbol{\theta}_q$ is a vector of coefficients. The residual υ_{iq} captures other unobservable factors and satisfies $E(\upsilon_{iq}|\boldsymbol{Z}_i) = 0$. Benefit receipt is modelled by a probit specification (benefit receipt model):

$$\tilde{R}_i = \boldsymbol{\beta} \boldsymbol{Z}_i + \gamma_1 \eta_{i1} + \cdots + \gamma_Q \eta_{iQ} + u_i \tag{4}$$

where the observed benefit receipt status $R_i = 1$ when $\tilde{R}_i > 0$ and $R_i = 0$ otherwise; $\boldsymbol{\beta}$ and the γ_q are coefficients and u_i^s is a stochastic disturbance term. While allowing correlation between the Q latent constructs, we make the standard assumption underlying probit models that the stochastic residual u_i is independent of $(\boldsymbol{Z}_i, \boldsymbol{\eta}_i)$ and the residuals in the measurement equations (1). In writing (3) and (4), we allow the same covariates to represent the influences on disability and on benefit claim behaviour. This is not necessary, and there may be exclusion restrictions (which are not necessary for identification) on the vectors $\boldsymbol{\beta}$ and $\boldsymbol{\theta}_q$.

We say that survey s is ***complete*** if the $J \times Q$ loadings matrix $\{\lambda_{jq}^s\}$ is of full column rank Q; this requires that, for each dimension of disability q, at least one of the j observed indicators D_{ij}^s has a non-zero loading λ_{jq}^s. In the Online

Appendix, we show that completeness is sufficient to identify the model under our assumptions. The surveys are said to be ***compatible*** if the assumption of common parameters across surveys in equations (3) and (4) is valid.

Several studies have shown that, in the older population, women tend to report significantly higher rates of functional difficulties than comparable men (Rahman and Liu 2000, Crimmins *et al.* 2011). Some researchers have attributed this apparent female functional disadvantage to higher true prevalence of nonfatal but disabling conditions such as arthritis and osteoporosis (Wingard 1984, Verbrugge and Wingard 1987). Others have found that, even when controlling for chronic conditions, women still report higher mean levels of functional disability (Waltz and Badura 1984). This could be due to a higher propensity for women to report ill health than men with the same underlying true health status (Verbrugge 1980, Hibbard and Pope 1983); or to heightened sensitivity to symptoms because of gender-specific social expectations and life experience (Verbrugge and Wingard 1987); or to task specificity if women are more engaged than men in household tasks that require actions such as bending and lifting. This measurement issue has been termed variously: 'state-dependent reporting bias' (Kerkhofs and Lindeboom 1995), 'scale of reference bias' (Groot 2000) and 'response category cut-point shift' (Lindeboom and van Doorslaer 2004). However, unless we can specify *a priori* a subset of indicators in each survey for which response behaviour is gender-invariant, it is impossible to distinguish the causal effect of gender on true latent disability from its effect on reporting behaviour. We allow for the possibility of inherent gender differences in disabilities by allowing the parameters of the measurement equations (1)-(2) to be gender-specific. We therefore exclude gender from equation (3).

We estimate the system comprising all equations (1)–(4) simultaneously allowing for the discrete nature of the dependent variables, using robust maximum likelihood as implemented in *MPlus* version 6.11 (Muthén and Muthén, 2010). This is done separately for each survey, to avoid imposing by assumption any homogeneity across surveys. All standard errors are clustered by household to allow for intra-household correlation. However, since we do not have access to indicators of the geographical primary sampling units used in the sampling designs in the FRS, we are not able to allow for geographical clustering, and the quoted standard errors are expected to understate sampling variation to a small extent. We have been able to confirm this for the ELSA sample, where psu and stratum indicators are available; standard errors increase to a negligible extent (details available on request). This suggests that the true size of our tests of between-survey parameter stability is very slightly larger than the nominal significance level, giving a small tendency to over-reject parameter stability, which increases the force of our eventual conclusions.

3 Data

The analysis is based on three sample surveys: the first wave of ELSA; the corresponding twelfth wave of BHPS; and the 2002/03 cross section of FRS. All three surveys have been widely used for research on physical health and disability: see, for example, Melzer *et al.* (2005), Banks *et al.* (2006), Mayhew *et al.* (2010), Chan *et al.* (2012) for ELSA; Benítez-Silva *et al.* (2009), Oswald and Powdthavee (2008), Banks *et al.* (2009) for BHPS; and Kasparova *et al.* (2007), Hancock and Pudney (2013) and Morciano *et al.* (2014) for FRS. Although the three surveys are broadly similar in sampling design, they differ considerably in their initial response, degree of cumulated attrition, and in methods of constructing weights intended to deal with departures from uniform sampling; see Table 1 for a summary of these differences.

The FRS has a sample size of over 25,000 private households. It is an annual cross-section and therefore suffers from nonresponse but not accumulated attrition. The FRS response rate in 2002/3 was 64% of eligible households (Campbell 2004). The BHPS started in 1991 and followed a sample of approximately 10,000 households annually, so our sample has come through twelve waves of attrition and possible panel conditioning. The initial BHPS response rate was 74% and 67% of those original respondents gave a full interview in wave 12 (Lynn *et al.*

2006). ELSA is a panel of individuals aged 50+ and their partners in approximately 8,000 private households in England. Panel membership is based on interview in the 1998, 1999 or 2001 Health Surveys for England (HSE). The wave 1 ELSA data are thus potentially affected by nonresponse in the HSE and a further round of attrition; HSE response rates were 74% (1998), 76% (1999) and 74% (2001) and of those selected for ELSA, around 70% responded to its first wave (Taylor *et al.* 2003). We choose the first wave of ELSA as our common time point to avoid the effects of subsequent attrition. We limit our analysis to people aged 65 years or over, living in England. The former restriction is because only people aged 65 or over can claim AA. The latter is imposed by the ELSA sampling frame. We also exclude respondents receiving Disability Living Allowance (a similar benefit that can be claimed before age 65) because DLA recipients cannot also claim AA.

The three surveys also differ in questionnaire content. The FRS collects very detailed income and benefit information, used as the basis for most official statistics on welfare and disability program targeting, but a limited set of disability indicators. ELSA provides a richer range of health and disability measures but slightly more limited income data than the FRS (for example, ELSA collects some income components gross of tax and others net). In the BHPS, it is not always possible to distinguish whether a particular income source is gross or net. BHPS information on health and disability is more detailed than the FRS in

Table 1: Comparing the FRS, ELSA and BHPS along sample design and structure, data collection and weighting procedures†

	FRS (2002/03)	ELSA (wave 1)	BHPS (wave 12)
Population coverage	People in private dwellings, UK	People in private dwellings, England	People in private dwellings, Great Britain
Timing	Cross-section, Apr 2002-Mar 2003	Longitudinal study, Mar 2002-Mar 2003	Longitudinal study, Sep 2002-Dec 2002
Frame	Royal Mail's small users's Postcode Address File (PAF)	1998, 1999 and 2001 Health Survey for England (HSE) samples drawn from different vintages of PAF. ELSA includes households with an adult of 50 or older who agreed to re-contact	PAF
Sample design	Sample design is an equal probability selection mechanism (EPSEM), with two-stage stratified random sampling	Two-stage stratified EPSEM design in the HSE	Two-stage stratified EPSEM design at wave 1 (1991)
Stratification variables	Region, socio-economic group profile, adult economic activity rate, male unemployment rate	Health Authority, proportion of households with a head of household in a non-manual occupation	Region, socio economic group profile, proportion of pensionable age, proportion of employed persons working in agriculture
Response rate	64%	HSE response rate 69%; 92% consent rate; 70% response rate at ELSA wave 1, giving 44% response overall	74% at wave 1; 50% allowing for cumulated attrition to wave 12
Weighting	*Design weights* adjust for selection of households within addresses. *Nonresponse weighting* is not used. *Calibration weights* are based on age, gender, lone parents/all families with children, housing tenure and Council Tax Band distributions from official statistics	*Nonresponse weights* compensate for unit nonresponse at HSE, refusal post-HSE and nonresponse in ELSA wave. ELSA phase uses inverse response probability from a logistic regression on age of the oldest household member, Regional Health Authority, household size, social class, year of HSE interview and longstanding illness, as observed in HSE datasets. *Calibration weights* match age-sex cell frequencies from the non-institutionalised population of the 2001 Census	*Design weights* adjust for selection of households within addresses. *Nonresponse weights* at household level based on region, socio-economic group and type of accommodation. At individual level, inverse response probability from logistic regression on region, housing tenure, affluence, household size, marital and employment status, age, sex. *Calibration weights* use 1991 and 2001 Census marginal distributions for household tenure, household size, no. of cars, age and sex
Question Wording on AA receipt	And looking at this card, are you at present receiving any of the state benefits shown on this card - either in your own right or on behalf of someone else in your household? [Attendance Allowance]	Have you/you or your husband/wife/partner received any of these health or disability benefits in the last year? [Attendance Allowance] Which of these health or disability benefits have you received in the last year? [Attendance Allowance] Which of these health or disability benefits are you receiving at the moment? [Attendance Allowance]	I am going to show you four cards listing different types of income and payments. Please look at this card and tell me if, since September 1st 2001, you have received any of the types of income or payments shown, either just yourself or jointly? [Attendance Allowance]

† Source: Campbell (2004), Taylor *et al.* (2003), Taylor *et al.* (2006), Taylor *et al.* (2007), Lound and Broad (2013).

some respects but less so than ELSA. The surveys differ in the information they collect by proxy for participants who are not able to provide responses themselves, in particular FRS collects information on disability and AA receipt from proxy respondents, whereas BHPS and ELSA do not. We return to treatment of proxy respondents below. Campbell (2004), Taylor *et al.* (2003) and Taylor *et al.* (2006) respectively give detailed descriptions of FRS, ELSA and BHPS sample design and data collection procedures.

In each survey, information about receipt of AA, recorded by the binary variable R_i, is collected through questions following those on health and disability. Thus, none of the three surveys is especially vulnerable to the justification bias in disability measurement that is a concern when the benefits module precedes the health module within the questionnaire (Crossley and Kennedy 2002). There are differences in the reference period for questions on AA receipt: the BHPS covers the year preceding interview; the FRS refers specifically to the time of interview; and ELSA asks separately about different reference points. For ELSA we use receipt of AA at the time of interview, to give comparability with the FRS.

A wide range of disability indicators is available in one or more of the three surveys. In this study, we use subjective indicators which are the most widely available in social surveys. Appendix Table A1 reports the functional limitation indicators D_j offered by each survey and used in our analysis, with their prevalence rates among AA recipients and non-recipients. Binary indicators in the FRS cover difficulties in eight areas of life. ELSA provides a longer list of indicators including limitations to specific Activities of Daily Living (ADL) (Katz *et al.*, 1963) or Instrumental Activities of Daily Living (IADLs) (Lawton and Brody 1969). The BHPS indicators include binary variables representing activities limited by health and a set of 6-point categorical variables, built from two questions on whether the respondent usually manages to perform a set of mobility and personal care activities alone or only with assistance, and whether he/she finds it very easy, fairly easy, fairly difficult or very difficult. There is a considerably higher sample prevalence of reported functional limitations among AA recipients than non recipients, consistently across surveys and specific indicators.

The choice of other personal characteristics included in $\boldsymbol{Z}$ is governed by previous work on the socio-economic gradient in health or disability (e.g. Goldman 2001) and on older people's benefit claim behaviour (for example Zantomio 2013 in relation to AA; Hernandez *et al.* (2007) and Pudney *et al.* (2006) for means-tested benefits). We use age (in the form of a spline with a knot at the median age across all samples of 73 to allow for non-linearity in the age gradient of disability and receipt of AA), gender, being educated beyond the compulsory minimum, housing tenure, and log equivalised pre-benefit income in both equations. Information on past occupation is not collected from pensioners in the FRS and therefore are not included in $\boldsymbol{Z}$. Income represents both the socio-economic gradient in health and the basic need for financial support which underlies benefit claim behaviour. It is derived as the sum of income from pensions, earnings, savings and other sources received by any member of the benefit unit (defined as an adult plus their spouse (if applicable) plus any dependent children they are living with), but excludes disability and means tested benefits. Disability benefits must be excluded from the latent disability equation because they are a consequence, and not a cause, of disability, and from the AA equation as it is income in the absence of AA that influences the decision to claim. Means-tested benefits are excluded because their level can also depend on disability through the Severe Disability Premium, an addition to the income thresholds used to assess entitlement to means-tested welfare benefits and applies where the claimant is receiving AA . To account for differences in benefit unit size we apply the modified OECD equivalence scale to income. For this older population, our income measure is dominated by pension income, which is a good indicator of past labour market success, itself strongly related to lifestyle characteristics which have associated health implications. Thus estimates of the impact of income on disability should be interpreted in this wide sense. Log income is entered as a spline with a knot at the median log income level (log of £615.70 per month, 2002 prices). Our definition of housing tenure distinguishes those who own their homes outright from those who rent or are still repaying their mortgage. Outright home-ownership is used to capture an additional long-term socio-economic influence on health. It also allows for the lower financial need (lower housing costs) that outright owners have compared with those who face rent or mortgage costs, to influence their benefit claim behaviour. Current partnership status (married/cohabiting versus single) is also included as a covariate in the AA receipt equation since it has previously been found to affect benefit claim behaviour (Hernandez *et al.* 2007; Pudney *et al.* 2006).

All variables have been derived in a consistent manner as far as possible, although perfect comparability cannot be guaranteed (sample means and standard deviations for the socio-economic characteristics $\boldsymbol{Z}$ observed in each sample are given in Table O3 of the Online Appendix). There are some differences between surveys: for example, ELSA sample members are significantly younger and more educated than their BHPS and FRS counterparts; the proportion of outright homeowners is higher in ELSA and the BHPS than in the FRS; and the mean of (log) income is significantly higher in the BHPS than in ELSA and the FRS. FRS reports a higher rate of AA receipt (9.7%) than ELSA or BHPS (7.2%). Comparisons with administrative data are not straightforward because they include AA recipients in the care-home population. We estimate that of the over 65 non-care home population, excluding those who received DLA, between 12.7 and 13.8% received AA in 2002. This is based on DWP statistics on recipients of AA and

DLA which include, but do not separately distinguish, recipients in care homes, together with estimates from Comas-Herrera et al. (2010) on the numbers of over 65s resident in care homes and the proportions of them who receive public support with the care home fees and are therefore not eligible to receive AA. All three surveys therefore seem to under-represent AA recipients but FRS less so than ELSA or BHPS.

Ideally we would use all proxy cases since they are likely to include some of the most severely disabled respondents. This view is supported by an analysis of proxy respondents in the FRS, revealing AA receipt to be about twice as high among proxy respondents as non-proxy respondents (18.1% against 9.1%). However we are forced to exclude proxy responses in ELSA (1.9%) and BHPS (4.1%) as their proxy questionnaires do not collect the respondent's disability (ELSA) or AA receipt (BHPS). We retain the larger proportion of proxy cases (6.5%) in the FRS which does collect this and other relevant information for proxy cases, using a proxy response as an additional disability indicator in the measurement model. After these exclusions and dropping cases with missing values for variables used in the analysis, the sample sizes are 1,042, 5,142 and 6,744 individuals from the BHPS, ELSA and FRS respectively. We also assess the sensitivity of the results to the exclusion of FRS proxy cases in which case the FRS sample is reduced to 6,308.

In the next section, we present results based on the full unweighted samples, and return to the issue of sample comparability in section 5.

4 Estimation results

4.1 The measurement model

To implement the model, we must specify the dimensionality of latent disability and choose a normalisation to deal with its non-observability and lack of natural units of measurement. Our main results come from survey-specific SEMs with a single latent disability factor and a simple normalisation. For the latter, we choose a priori one indicator from each survey that appears to be based on essentially the same question. These are: the FRS question about mobility ('moving about'); the ELSA question about capacity to 'walk 100 yards'; and the BHPS question about 'walking more than 10 minutes'. We then normalise the factor loading for each of these indicators to be unity. In section 6, we explore the sensitivity of the results to our choice of normalisation and number of factors. Controversy exists over whether functional disability should be treated as a one dimensional or multi-dimensional construct (see for instance Fitzgerald et al. 1993; Spector and Fleishman 1998). As a check on the robustness of our main model, in section 6 we also estimate a 2-factor model distinguishing physical and cognitive disabilities. Although passing reference is often made to the multi-dimensional nature of disability, we are not aware of any previous estimates of multi-factor models of this kind in the existing literature.

The estimates of the measurement model are presented in Table 2: the factor loadings λ^s_{jq}, representing the effect of latent disability η on each indicator D^s_{ij}, are positive and highly significant in each survey. Although the pattern of estimated factor loadings is similar for male and female respondents in each survey, there are significant differences. In FRS, the loading associated with 'lifting, carrying or moving objects' is significantly higher for women. In ELSA, factor loadings associated with reported difficulties in ADLs like 'bathing or showering', 'eating', 'getting in or out of bed' and 'using the toilet' and IADLs like 'doing work around the house or garden' are significantly lower for women; in BHPS, a significantly lower factor loading for women is also found for difficulties in bed transfers and 'bathing or showering'.

The Akaike information criterion suggests that the unrestricted models (which allow the parameters of the measurement equations (3) to be gender-specific) provide slightly better balances of model fit and parsimony. This result is also confirmed by the Satorra-Bentler (2001) test at the 1% level for each of the three surveys.

Table 2: Estimated 1-factor models

Disability Indicator§§	Factor loading	Standard error	Disability Indicator§§	Factor loading	Standard error
		MEN			
	FRS			*ELSA*	
MOBILITY	1	-	WALK100	1	-
LIFTING	1.005†	(0.088)	SITTING	0.386†	(0.031)
DEXTERITY	0.723†	(0.065)	CHAIR	0.581†	(0.040)
CONTINENCE	0.395†	(0.037)	CLIMBSEV	0.724†	(0.049)
COMMUNIC	0.385†	(0.042)	CLIMB1	0.990†	(0.066)
MEMORY	0.420†	(0.042)	STOOP	0.641†	(0.043)
DANGER	0.510†	(0.093)	ARMS	0.503†	(0.042)
OTHER	0.098†	(0.027)	PULL/PUSH	1.008†	(0.078)
PROXY	0.116†	(0.029)	LIFTING	0.934†	(0.066)
	BHPS		COIN	0.379†	(0.047)
HOUSEWORK	0.851†	(0.126)	DRESSING	0.661†	(0.048)
STAIRS	0.959†	(0.129)	WALKING	1.052†	(0.134)
DRESS	0.660†	(0.114)	BATH	0.863†	(0.068)
WALKING	1	-	EATING	0.596†	(0.087)
STAIRS	1.112†	(0.180)	BED	0.879†	(0.085)
MOBILITY	1.358†	(0.275)	TOILET	0.738†	(0.091)
BED	1.346†	(0.259)	CONTINENCE	0.299†	(0.030)
NAILS	0.585†	(0.085)	MAP	0.406†	(0.049)
BATH	1.001†	(0.171)	MEAL	0.806†	(0.101)
ROAD	1.151†	(0.176)	SHOPPING	1.018†	(0.084)
			PHONE	0.358†	(0.046)
			MEDICATION	0.477†	(0.071)
			HOUSEWORK	1.132†	(0.086)
			MONEY	0.453†	(0.057)
		WOMEN			
	FRS			*ELSA*	
MOBILITY	1	-	WALK100	1	-
LIFTING	1.186†	(0.102)	SITTING	0.399†	(0.029)
DEXTERITY	0.643†	(0.047)	CHAIR	0.532†	(0.033)
CONTINENCE	0.431†	(0.035)	CLIMBSEV	0.671†	(0.043)
COMMUNIC	0.365†	(0.037)	CLIMB1	0.899†	(0.053)
MEMORY	0.416†	(0.036)	STOOP	0.653†	(0.040)
DANGER	0.426†	(0.052)	ARMS	0.500†	(0.035)
OTHER	0.060‡	(0.024)	PULL/PUSH	0.899†	(0.056)
PROXY	0.121†	(0.024)	LIFTING	0.900†	(0.058)
	BHPS		COIN	0.433†	(0.037)
HOUSEWORK	0.968†	(0.149)	DRESSING	0.650†	(0.042)
STAIRS	1.201†	(0.168)	WALKING	0.959†	(0.090)
DRESS	0.910†	(0.167)	BATH	0.722†	(0.047)
WALKING	1	-	EATING	0.428†	(0.055)
STAIRS	0.911†	(0.129)	BED	0.686†	(0.054)
MOBILITY	1.066†	(0.164)	TOILET	0.577†	(0.051)
BED	0.965†	(0.151)	CONTINENCE	0.251†	(0.022)
NAILS	0.582†	(0.080)	MAP	0.343†	(0.029)
BATH	0.777†	(0.112)	MEAL	0.811†	(0.074)
ROAD	1.110†	(0.163)	SHOPPING	1.135†	(0.080)
			PHONE	0.327†	(0.045)
			MEDICATION	0.479†	(0.073)
			HOUSEWORK	0.926†	(0.061)
			MONEY	0.479†	(0.048)

Statistical significance: † $p < 0.01$; ‡ $p < 0.05$; § $p < 0.1$. §§ A more detailed description for each D_j^s indicator can be found in Online Appendix Table O1.

4.2 The disability model

Estimates for the model (3) of latent disability status are reported in Table 3, together with t-tests of individual coefficient equality and the overall χ^2 Wald tests for equality of the whole coefficient vector for each pair of surveys. The conditional mean of latent disability η increases with age: the FRS and ELSA display a nonlinear relation between age and disability, with a higher gradient beyond age 73. In the BHPS we find a strong and near-linear relationship between age and disability. Higher education and pre-benefit income are associated with lower disability, giving evidence of a socio-economic gradient in disability that is consistent across surveys. Being a homeowner decreases the conditional mean of η, particularly in ELSA. The variance of the latent disability factor is greater in the BHPS than in the FRS or ELSA, but we find that the factor variances are quite comparable across surveys (a 10% significant difference is found only for the FRS-ELSA contrast). The estimated coefficients for FRS and ELSA are similar in size and the Wald test cannot reject the hypothesis of equality; when the BHPS is used as the basis for comparison, the null hypothesis of joint equality of coefficients is rejected (P-values 0.064 and 0.028).

Table 3: Estimates of the latent disability equation

Covariates	Coefficients			Tests and coefficient differences		
	FRS	ELSA	BHPS	FRS-ELSA	FRS-BHPS	ELSA-BHPS
Spline age 65-73	0.038†	0.035†	0.127†	0.003	-0.089†	-0.092†
	(0.013)	(0.012)	(0.036)	(0.018)	(0.038)	(0.038)
Spline from age 73+	0.091†	0.099†	0.128†	-0.008	-0.037§	-0.029
	(0.008)	(0.008)	(0.020)	(0.011)	(0.022)	(0.022)
Post-compulsory education	-0.279†	-0.28†	-0.182	0.001	-0.096	-0.097
	(0.065)	(0.061)	(0.149)	(0.089)	(0.163)	(0.161)
Income spline to median	-0.162†	-0.046	-0.172§	-0.116§	0.009	0.125
	(0.047)	(0.052)	(0.104)	(0.070)	(0.114)	(0.116)
Income spline from median	-0.336†	-0.310†	-0.558†	-0.025	0.223	0.248
	(0.085)	(0.072)	(0.206)	(0.111)	(0.223)	(0.218)
Outright owner	-0.382†	-0.487†	-0.185	0.105	-0.197	-0.302§
	(0.064)	(0.064)	(0.151)	(0.090)	(0.164)	(0.163)
	3.012†	2.543†	3.298†	0.469§	-0.286	-0.755
Variance (σ_v^2)	(0.275)	(0.225)	(0.788)	(1.320)	(0.343)	(0.921)
	Sample size			*Coefficient equality* $^{2}(6)$		
	6,744	5,142	1,042	4.361	11.920§	14.139‡

Statistical significance of the coefficient, t-test cross-sample coefficient difference and χ^2 statistic: † p < 0.01; ‡ p < 0.05; § p < 0.1. Standard Errors in parentheses.

4.3 The benefit receipt model

Estimates for equation (4), describing the relationship of AA receipt with socio-economic characteristics and latent disability, are reported in Table 4. Receipt of AA is clearly disability-related in each of the surveys, and disability consistently emerges as the dominant variable in explaining AA receipt. Although disability might raise barriers to claiming and at the same time reduce individuals' capacity to benefit from additional cash income, the survey evidence suggests there is successful targeting of AA on the disabled older population, irrespective of the

Table 4: Estimates of the equation for receipt of Attendance Allowance

	Coefficients			Coefficient differences		
Covariates	FRS	ELSA	BHPS	FRS-ELSA	FRS-BHPS	ELSA-BHPS
Latent disability η	0.569†	0.477†	0.538†	0.092§	0.031	-0.060
	(0.041)	(0.035)	(0.095)	(0.054)	(0.103)	(0.101)
Female	0.122§	0.251†	-0.068	-0.129	0.190	0.319§
	(0.065)	(0.073)	(0.172)	(0.098)	(0.184)	(0.187)
Spline age 65-73	-0.040†	-0.036†	-0.084†	-0.004	0.043§	0.048‡
	(0.008)	(0.007)	(0.021)	(0.011)	(0.022)	(0.022)
Spline from age 73+	0.058†	0.046†	0.028§	0.012	0.030§	0.017
	(0.006)	(0.007)	(0.015)	(0.009)	(0.016)	(0.016)
Post- compulsory education	-0.161‡	-0.238†	-0.070	0.077	-0.090	-0.167
	(0.065)	(0.071)	(0.155)	(0.096)	(0.168)	(0.171)
(ln) income spline to median	-0.008	-0.092§	-0.041	0.083	0.033	-0.050
	(0.048)	(0.049)	(0.090)	(0.069)	(0.102)	(0.102)
(ln) income spline from median	-0.392†	-0.422†	-0.411§	0.030	0.019	-0.011
	(0.120)	(0.154)	(0.247)	(0.195)	(0.274)	(0.291)
Outright owner	-0.136‡	-0.006	-0.265	-0.130	0.128	0.259
	(0.062)	(0.071)	(0.164)	(0.095)	(0.175)	(0.178)
Married/cohabiting	-0.076	0.087	-0.171	-0.163	0.094	0.257
	(0.064)	(0.076)	(0.182)	(0.100)	(0.193)	(0.198)
$\chi^2(9)$ test of coefficient equality				14.398	14.685§	14.844§

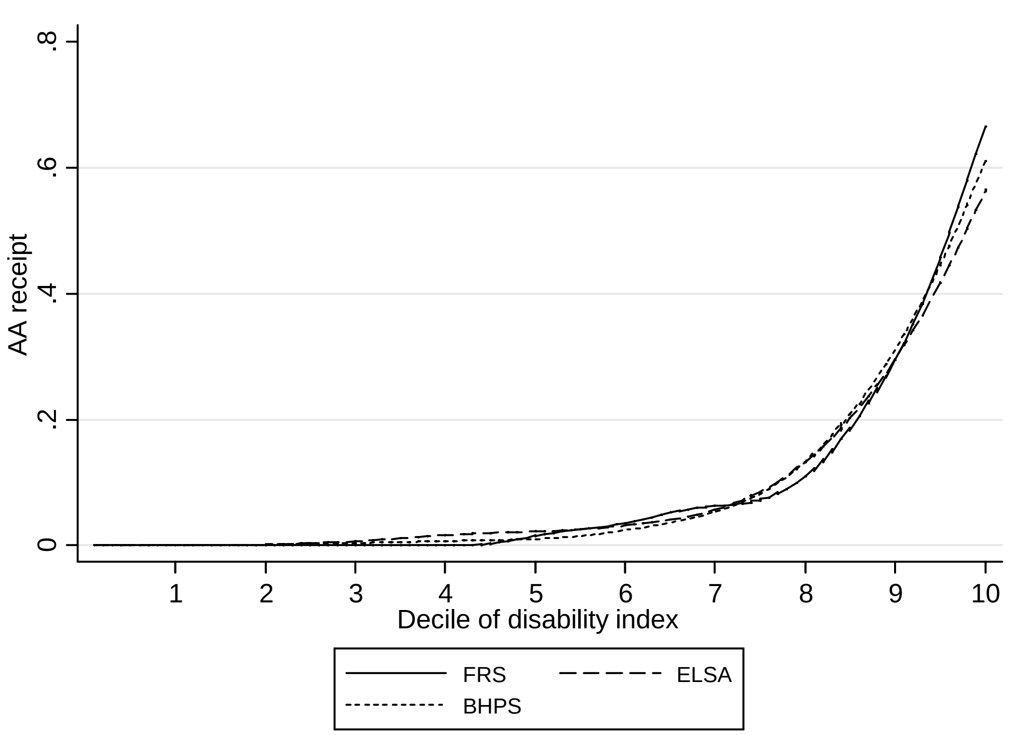
Statistical significance of the coefficient, t-test cross-sample coefficient difference and χ^2 statistic: † $p < 0.01$; ‡ $p < 0.05$; § $p < 0.1$. Standard Errors in parentheses.

source of survey data. This is clear from Figure 1, which shows the mean prevalence of AA receipt within each decile of the distribution of the posterior prediction of latent disability for each individual. The strong disability-targeting of AA emerges very clearly for all three surveys.

The estimated probability of receiving AA declines nonlinearly with income. We find that, below median income, the coefficient is significant at the 10% level only in ELSA, so the income gradient in AA receipt operates primarily among higher-income people. The negative gradient is due both to the low incidence of disability among high-income groups (Pudney 2010) and to the low propensity of these groups to claim benefit (Hernandez *et al.* 2007). Consequently, although AA is not means-tested, patterns of receipt mimic to some degree the effect of means testing for those in the top half of the pensioner income distribution.

We find significant evidence of a negative association between the level of education and AA receipt in both ELSA and FRS. This suggests that any advantage that more educated people may have in navigating the benefits system is outweighed by factors such as less contact throughout their lives with the benefit system, or greater perceived stigma from claiming benefits (as also found in Zantomio 2013). Owning one's home outright reduces significantly the probability of AA receipt in the FRS and the BHPS. This could reflect a lower financial need among homeowners, or the same factors that may be at work for more educated people could play a similar role for outright homeowners.

Figure 1: Proportion of people in receipt of AA by predicted severity of disability



Note: Smoothed local linear regressions applied on the FRS (solid line), the ELSA (long dashed line) and the BHPS (dotted line) samples. Bandwidth set equal to 0.4.

Receipt of AA appears gender-related in the FRS and the ELSA, where men are less likely to receive AA than women; gender differences are insignificant in the BHPS. In all three surveys, age affects the probability of AA receipt non-linearly, with a convex age profile. There is again a significant difference between the estimated age profile for the BHPS compared with FRS and ELSA, with a less significant upturn at older ages. Finally, none of the surveys suggests that the presence of a partner significantly affects the probability of receiving AA. Inspection of coefficients in this piecemeal way creates a bias in favour of finding significant differences, because of the multiple comparisons involved. However, a joint Wald test finds a significant difference between BHPS and the other two samples (P-values 0.100 and 0.095). We do not reject coefficient equality between the FRS and ELSA.

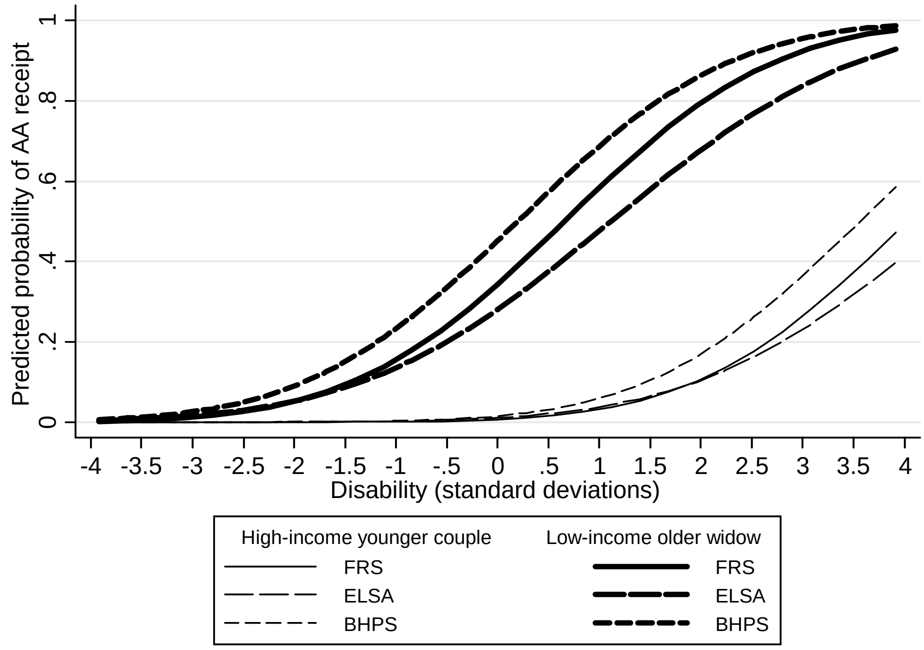
In Figure 2(a), we compare the implications of the estimated models, for two illustrative individuals: a 65-year old man living with his partner as an outright homeowner with income 50% above the median; and an 85-year old non-home-owner widow, with equivalised income 75% of the median. Both have compulsory minimum education. In Figure 2a, the between-survey differences in their AA-disability profiles are modest in comparison with the predicted differences between hypothetical individual types. For example, at a disability level one standard deviation above the mean, the three models predict a 4-7% rate of receipt for the couple compared to a 50-71% rate for the widow. At disability level of 2.5 standard deviations above the mean, the ranges are 16-26% for the couple and 77-92% for the widow.

In Figure 2(b), we compare the estimated AA-income profiles. Again, the between-survey differences in these profiles are modest in comparison with the predicted differences between hypothetical individual types. The rate of receipt for the low-disability type (at the 25th percentile of the disability index distribution) couple is essentially zero, whereas the rate of receipt for the high-disability type (at the 75th percentile of the disability index distribution) ranges from 31% to 37% in the income interval we consider. The rate of receipt is nonlinear in income: almost flat below median equivalised income and steadily declining thereafter. For example, the rate of receipt for the highly-disabled widow ranges from 34 to 39% at the 25th (£435 per month) and at the 50th percentile of the income distribution, and 27-33% at the 75th percentile (£917 per month).

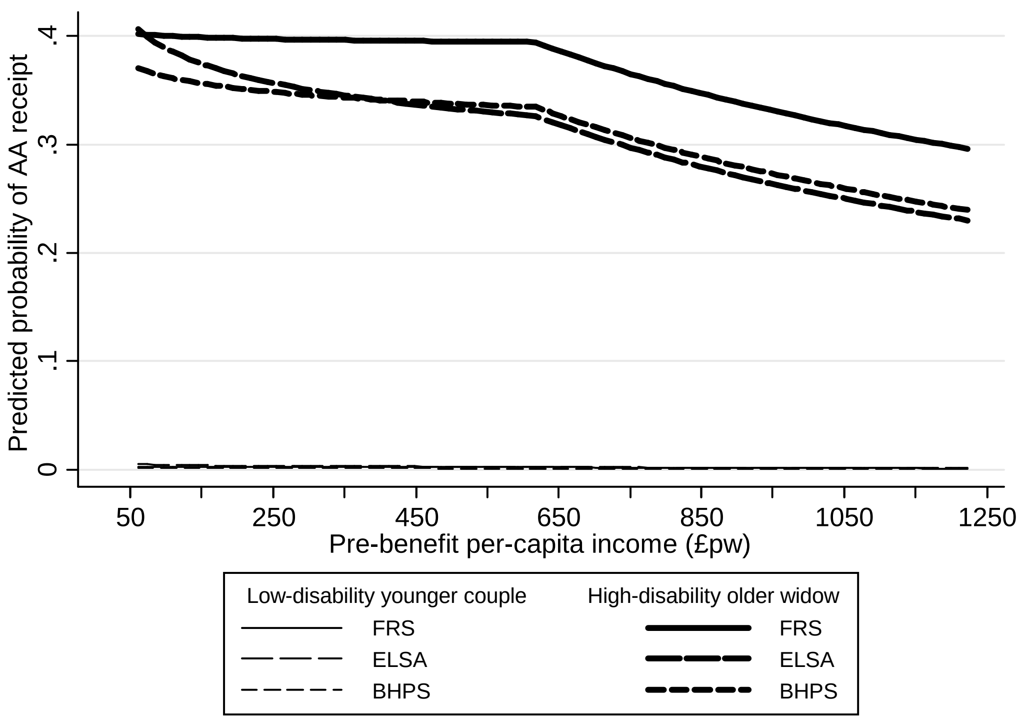
In general, the three surveys show similar patterns in terms of their empirical AA-disability relationship. However at some disability levels between survey differences in predicted probabilities of AA receipt are sizeable. The between-survey differences are statistically significant when the BHPS is used as the basis for comparison. In the next section we investigate the extent to which these differences might be attributable to differences in sample composition.

Figure 2: Predicted probabilities of AA receipt by survey for two benchmark cases

(a) The AA-disability relation



(b) The AA-income relation



5 Controlling sample composition

If statistical models are empirical approximations local to the region spanned by the sample data, then cross-survey differences in model estimates might result only from differences in their covariate distributions rather than any more fundamental measurement problem. As Table 1 makes clear, there are important differences in the empirical distribution of the covariates in the three surveys, resulting from the differences in design and patterns of response.

In single-survey analysis, the standard method of controlling sample composition is to use survey weights. Broadly, these have three elements: *design weights* which compensate for deliberate non-uniform sampling rates across the population; *nonresponse weights* which compensate for variations in response probabilities across individuals and households with different characteristics; and *calibration/ post − stratification weights* used as a final step to bring the sample composition in line with whatever is known about the structure of the population. If the assumptions underlying the derivation of weights (e.g. missingness at random (MAR)) are valid and if the weights are implemented in the "correct" way by each survey, then separate weighted samples should identify essentially the same population parameters, if the questionnaires have the same informational content. However, the weighting strategies are not harmonised across the three surveys (see Table 1 for a summary of the weighting procedures). Different covariates appear in response models used to generate nonresponse weights and the calibration stage is done in different ways. Given these methodological conflicts, it is unlikely that the use of the weights supplied with each survey will solve the comparability problem, and it is even possible for weighting to impair, rather than improve, comparability. Nevertheless, we have carried out weighted analyses and found that weighting does not fully eliminate the between-survey differences we found in section 4. For the disability equations, the Wald χ^2 of coefficient equality P-value slightly rises from 0.064 to 0.102 for the FRS-BHPS comparison and it decreases from 0.028 to 0.026 for the ELSA-BHPS comparison (see Appendix Table A2). For the Attendance Allowance equation, the Wald χ^2 P-value decreases from 0.100 to 0.048 for the FRS-BHPS comparison and it rises from 0.095 to 0.142 for the ELSA-BHPS comparison (see Appendix Table A3).

Matching techniques provide another way of reducing bias from differences in the sampling distribution of covariates across surveys. They involve estimating the models using survey-specific subsamples which are balanced in terms of the set of common covariates thought to influence disability and AA receipt. The matching approach has not been widely used in this context, but there are some precedents (Rosenbaum, 2002; D'Orazio et al., 2006; Rässler, 2002). The method requires (at least partial) common support across surveys for the matching variables, which holds in our samples (see Table O2 of the Online Appendix). We make the assumption that the matching variables are comprehensive in the sense that, conditional on them, sub-sample selection can be regarded as random. This is essentially the same MAR assumption underlying weighting methods and, although untestable, is plausible, given the three surveys' sample design.

In practice, we take each survey in turn as a baseline and construct matched sub-samples from the other two surveys, yielding six pairs of matched samples. The matching algorithm (Leuven and Sianesi 2003) uses one-to-one nearest-neighbour matching, minimising the Mahalanobis distance for the variables age, gender, post-compulsory education, partnership, housing tenure and log pre-benefit net income. Matching is performed without replacement, to avoid repeated use of the same observation from the matched survey, at the cost of possibly reducing the size of successfully matched samples. According to available sample size, in each round of pairwise matching we impose a caliper (ranging from 0.04 to 0.5) to prevent poor matches, equivalent in practice to exact matching of binary variables and very close matching for the continuous income and age variables; t-tests for the equality of means between each baseline sample and the corresponding matched samples were used to confirm the success of the algorithm in balancing the conditioning covariates. We also discarded matched pairs of observations whose income difference was in the top 5% when matching BHPS to ELSA and the top

10% when matching ELSA to BHPS. Means of socio-economic variables and AA receipt in the matched samples are given in Table O2 of the Online Appendix.

We repeated estimation of the system of equations (1), (3) and (4) on each of the six pairs of matched samples. Results obtained for the measurement equations (1)-(2) confirm the patterns described in Section 4, with mobility indicators playing a dominant role as indicators of latent disability. The three panels of Appendix Table A4 report estimated regression coefficients for the latent disability equation (3) obtained from samples mimicking the FRS, ELSA and BHPS sample compositions respectively. As in the unmatched samples (Table 3), we obtain significant disability gradients in age (positive) and income (negative) consistently across surveys, although some coefficients lose significance in smaller samples. Using separate t-tests of cross-sample coefficient stability, we would reject the null hypothesis of coefficient equality only for the first spline of income coefficient (at the nominal 5% level), when FRS or ELSA are used to mimic the BHPS sample composition. However, none of the individual t-tests would be significant if a Bonferroni correction were used, and the striking similarity of estimated coefficients is confirmed by the χ^2 tests of coefficients' joint equality: in none of the six paired survey comparisons is the null hypothesis rejected.

Estimated coefficients for the AA receipt equation (4) are reported in Appendix Table 5. The positive disability gradient in AA receipt found in the unmatched samples (Table 4) is also evident in the matched samples: estimates for the disability coefficient γ are positive, significant and remarkably similar in size. The negative income gradient is also confirmed, except for an insignificant positive coefficient when ELSA mimics the BHPS sample composition. The negative association between homeownership and receipt of AA is again found whenever the coefficient on homeownership is significant. For age, coefficient equality is rejected at the 5% level only for the second spline when BHPS observations are used to mimic the ELSA sample composition; but such isolated rejections are likely to arise from sampling error when large numbers of individual t-tests are used, and none would be significant if a Bonferroni correction were used. Joint Wald χ^2 tests of coefficient equality again fail to reject the hypothesis of coefficient equality in any of the six pairwise comparisons.

6 Robustness

6.1 The number of factors

In the estimated 1-factor measurement models of Table 2, there is a strikingly low correlation between the latent disability index and those indicators which might be thought to represent cognitive rather than physical disability. To allow for a distinction between physical and cognitive disability, we have also estimated a 2-factor model for each sample, following an exploratory factor analysis of the disability indicators. The attempt failed for the BHPS, where only a single factor could be detected, arguably because the BHPS disability questions lack completeness and have poor sensitivity to the cognitive dimension of disability. For the FRS and ELSA 2-factor models can be estimated (see Appendix Tables A9-A10 and O3 of the Online Appendix). The second factor appears to distinguish the cognitive aspect of disability for the FRS where difficulties in communication, in memory/concentration/ learning/understanding and in recognising physical danger are fairly obviously related to cognitive functioning. Since incontinence could stem from physical and/or cognitive problems, we allow for a cross-loading between the 2 factors for difficulties with continence. In ELSA, the second factor is determined from four cognitively-demanding IADLs (using a map, telephone use, self-medication, and handling finances) and, as for the FRS, we allow a cross-loading for continence. It is well known that there are limitations in the extent to which IADLs capture difficulties in cognitive functioning (Cromwell *et al.* 2003). We find the two factors to be strongly correlated (a similar result for the US is reported by Wallace and Herzog 1995). In the 2-factor latent disability equations (Table A6) the estimated coefficients for the first factor are close to those found in the 1-factor model for ELSA but are generally lower for the FRS, particularly for age and home-ownership. Using unmatched samples, we can reject the hypothesis of equal coefficients in the FRS and ELSA models for latent disability factor 1 but not factor 2 (Table A6). Results in Table A7 suggest a larger role for physical than cognitive influences on AA receipt with statistically insignificant differences between the estimated coefficients in the two surveys (P-values 0.140 and 0.192, respectively). The 2-factor specification confirms our previous findings on the relationship of AA receipt to socio-economic characteristics, since tests of coefficient equality do not reject the null hypothesis that coefficients ($\boldsymbol{\beta}$) of the observed covariates in the 2-factor models are equal to those obtained with the 1-factor specification in both surveys. The estimated coefficients of the 2-factor models are similar in size for FRS and ELSA. Based on a Wald-test, we reject the hypothesis of equality for the full AA coefficient vector $(\boldsymbol{\beta}, \boldsymbol{\gamma})$ (P-value $= 0.013$) but we do not reject for $\boldsymbol{\beta}$ alone (Wald P-value $= 0.244$). Cross-survey differences in the magnitude of the coefficients are not large and, for practical research purposes, one would draw essentially the same conclusions from the FRS and ELSA results.

6.2 Alternative normalisations

The 1-factor models set out above were estimated under the normalisation to unity of the factor loading associated with difficulties in mobility in each survey. Here we discuss the robustness of those findings to two alternative normalizations of : in the first, we constrain an alternative factor loading; in the second, we set the residual variance of η equal to 1.

The comparability of estimates of the disability and AA equations can be improved by normalising the loadings of more similar questionnaire items. For instance, the FRS and ELSA have questions on the capacity to lift weights (variable LIFTING) which are arguably more similar than those on general mobility. When the factor loading for LIFTING is normalised to unity, the concordance between the FRS and ELSA disability equation and AA coefficients does indeed improve, with the Wald χ^2 P-values rising to 0.271 and 0.287 respectively (1-factor specification, unmatched samples). Details of the estimates are in the Online Appendix, Tables O4-O6. However, the scope of this exercise is limited by the lack of a directly comparable indicator in the BHPS.

6.3 Proxy cases in the FRS

Since we are forced to exclude proxy cases from the analysis of ELSA and BHPS, we investigate the consequences of also excluding them from the FRS and dropping the proxy indicator from the disability measurement equations (see Tables O7-O9 of the Online Appendix). This has the effect of changing slightly the factor loadings on the other indicators. Nevertheless, all factor loadings remain positive and highly significant. The largest changes in loadings are for men, where the factor loading on lifting increases from 1.005 to 1.039, while those for memory problems and recognising when in danger fall from 0.420 to 0.356 and from 0.510 to 0.355 respectively. The estimated latent disability and AA receipt equations are not changed substantially. However, there are some small effects on the statistical significance of differences between the surveys in the estimated coefficients.

In both the disability and the receipt of AA equations, after dropping proxy cases, the differences between the FRS and ELSA become smaller but increase slightly when FRS is contrasted with BHPS.

7 Conclusions

Our aim in this study is to contribute to the current policy debate over reform prospects for the social care system by investigating the robustness of survey-based evidence on the targeting of public support for older people with disabilities. We have examined the three UK surveys (FRS, ELSA and BHPS) which have been the basis for much of the empirical analysis underpinning the debate on policy on disability in the pensioner population. Despite differences between the three surveys in terms of their questionnaire content, we have found that they have a coherent story to tell about the targeting of one form of public support in relation to disability, income and other personal and household characteristics.

We also claim to offer some advance in terms of the statistical modeling methodology typically used in the disability research literature. Adopting a latent variable approach, we are able to exploit the existence of multiple – but largely arbitrary and individually unreliable – survey indicators, whilst avoiding the common practice of using ad hoc count indices as disability measures. Results confirm that the probability of receiving AA increases strongly with the severity of disability and decreases with income – especially for those in the top half of the income distribution – after allowing for the socio-economic gradient in health that associates higher living standards with lower disability. This is important in the context of renewed suggestions that consideration be given to means testing AA (Commission on the Future of Health and Social Care in England, 2014). Contrary to some suggestions, we can say there is no evidence of people receiving AA without any disability revealed by their survey interview. In allowing for two latent disability factors we find evidence from the FRS and ELSA that physical disability has a larger influence on AA receipt than cognitive disability. Limitations in the BHPS survey instrument meant that we were unable to confirm this in the BHPS. This suggests that survey designers should be concerned more to ensure that disability indicators capture a range of types of disability rather than with the merits of each individual indicator. Our use of Mahalanobis matching to improve comparability by removing differences in sample composition also provides a valuable reminder of the need to consider sample coverage as a factor when reviewing a range of research findings.

References

Banks, J., Kapteyn, A., Smith, J. P. and van Soest, A. (2009) Work Disability is a Pain in the ****, Especially in England, the Netherlands, and the United States. In: *Health at Older Ages: The Causes and Consequences of Declining Disability among the Elderly* (eds D. M. Cutler and D.A. Wise), pp. 251-293. National Bureau of Economic Research.

Banks, J., Marmot, M., Oldfield, Z. and J.P. and Smith, J.P. (2006) Disease and disadvantage in the United States and in England. *Journal of the American Medical Association*, **295**, 2037–2045.

Benítez-Silva, H., Disney, R. and Jiménez-Martín, S. (2009) Disability, capacity for work and the business cycle: an international perspective, *Fundacin de Estudios de Economa Aplicada Working Papers,* 2009-28.

Berthoud, R. and Hancock, R. M. (2008) Disability benefits and paying for care, in *Advancing Opportunity: Social Care* (ed N. Churchill). London: The Smith Institute.

Bollen, K. (1989) *Structural Equations with Latent Variables*, New York: Wiley.

Bound, J. (1991) Self-Reported versus Objective Measures of Health in Retirement Models. *Journal of Human Resources* **26**(1), 106-138.

Cabinet Office (2010) *State of the Nation Report: Poverty, Worklessness and Welfare Dependency in the UK*. London: Cabinet Office.

Campbell, A. (2004) *Family Resources Survey: Annual Technical Report: 2002/03*, London: Office for National Statistics. Available at: http://www.statistics.gov.uk/downloads/theme_social/FRS_Tech0203.pdf

Chan, K.S., Kasper, J.D. and Pezzin, L.E. (2012) Measurement equivalence in ADL and IADL difficulty across international surveys of aging: findings from the HRS, SHARE, and ELSA. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, **67**(1), 121-132.

Comas-Herrera, A., Wittenberg, R and Pickard, L. (2010) The long road to universalism? Recent Developments in the Financing of Long-term Care in England. *Social Policy and Administration,* **44** (4), 375-391

Commission on Funding Care and Support (2011) *Fairer care funding: the report of the commission on funding care and support*. Available at: https://www.wp.dh.gov.uk/carecommission/files/2011/07/Fairer-Care-Funding-Report.pdf

Commission on the Future of Health and Social Care in England (2014) A new settlement for health and social care: interim report. London: the King's Fund.

Crimmins, E. M., Ki Kim J. and Solé-Auró A. (2011) Gender differences in health: results from SHARE, ELSA and HRS. *European Journal of Public Health,* **21**(1), 81-91.

Cromwell, D. A., Eagar K. and Poulos R. G. (2003) The performance of instrumental activities of daily living scale in screening for cognitive impairment in elderly community residents. *Journal of Clinical Epidemiology*, **56**(2), 131-137

Department of Health (2009) *Shaping the Future of Care Together*. Available at: http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_102338

Department of Health (2013) *Policy statement on care and support funding reform and legislative requirements*. London: Department of Health. Available at: http://www.dh.gov.uk/health/files/2013/02/Policy-statement-on-funding-reform.pdf

Department for Work and Pensions (2013) *AA claim form*. Available at: http://www.dwp.gov.uk/advisers/claimforms/aa1a_print.pdf

D' Orazio, M., Di Zio, M. Scanu, M. (2006) *Statistical Matching: Theory and Practice*, New York: Wiley.

Fitzgerald, J. F., Smith, D. M., Martin, D. K., Freedman, J. A., and Wolinsky, F. D. (1993), Replication of the multidimensionality of activities of daily living, *Journal of Gerontology*, **48**(1), S28–S31.
Forder, J. and Fernandez, J-L. (2009) *Analysing the costs and benefits of social care funding arrangements in England: technical report*. London: LSE. PSSRU Discussion Paper 2644. Available at: http://eprints.lse.ac.uk/24977/

Goldman, N. (2001) Social inequalities in health: disentangling the underlying mechanisms. In *Strengthening the Dialogue between Epidemiology and Demography* (eds M. Weinstein and A. Hermalin), Annals of the New York Academy of Sciences.

Groot, W. (2000) Adaptation and scale reference bias in self assessments of quality of life. *Journal of Health Economics*, **19**, 403–420.

Hancock, R. M. and Pudney, S. E (2013). Assessing the distributional impact of reforms to disability benefits for older people in the UK: implications of alternative measures of income and disability costs. *Ageing and Society* **34**, 232-257. DOI: http://dx.doi.org/10.1017/S0144686X1200075X

Hernandez, M., Pudney, S. E. and Hancock, R. M. (2007) The welfare cost of means-testing: pensioner participation in Income Support. *J. Appl. Econometrics,* **22**, 581-598.

Hibbard, J. H. and Pope, C. R. (1983) Gender roles, illness orientation and use of medical services. *Social Science and Medicine,* **17**, 129-137.

Hirst, M. (2004) The British Household Panel Survey: a longitudinal perspective on informal care (eds S. Becker and A. Bryman) *Understanding Research for Social Policy and Practice*, pp.190-193. Bristol: The Policy Press.

Jagger C., Matthews R., King D., Comas-Herrera A., Grundy E., Stuchbury R, Morciano M., Hancock R. and the MAP2030 team (2009) *Calibrating disability measures across UK national surveys*. Report prepared for Department of Work and Pensions. Available at: http://www.esrc.ac.uk/my-esrc/grants/RES-339-25-0002/outputs/read/04254e47-90ff-4f77-af7a-0533490acd3c

Johnson, R.J. and Wolinsky F.D. (1993) The Structure of Health Status Among Older Adults: Disease, Disability, Functional Limitation, and Perceived Health. *Journal of Health and Social Behavior*, **34** (2), 105-121.

Kasparova D., Marsh A. and Wilkinson D. (2007) *The take-up rate of Disability Living Allowance and Attendance Allowance: feasibility study*. Research Report No 442. London: Department for Work and Pensions.

Katz S., Ford A.B., Moskowitz R.W., Jackson B.A., Jaffe M.W. and Cleveland M.A. (1963) Studies of Illness in the Aged. The Index of ADL: A Standardized Measure of Biological and Psychosocial Function. *J. Am. Medical Ass.* **185**(12), 914-919.

Karlsson M., Mayhew, L., Plumb, R. and Rickayzen B. (2006) Future cost for long term care: cost projections for long term care for older people in the United Kingdom. *Health Policy*, **75**, 187–213.

Lawton M.P. and Brody E.M. (1969) Assessment of older people: self-maintaining and instrumental activities of daily living. *The Gerontologist,* **9**(3), 179-186.

Lee L-F. (1982) Health and wage: A simultaneous equation model with multiple discrete indicators. *International Economic Review,* **23**, 199-221.

Leuven, E. and Sianesi B. (2003) *PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing*. Available at: http://ideas.repec.org/c/boc/bocode/s432001.html.

Lindeboom, M. and van Doorslaer, E. (2004) Cut-point shift and index shift in self-reported health. *Journal of Health Economics,* **23**, 1083–1099.

Lynn, P. (ed.), with Buck, N., Burton, J., Laurie, H. and Uhrig S.C.N. (2006) *Quality Profile: British Household Panel Survey: Waves 1 to 13, 1991-2003.* Institute for Social and Economic Research, University of Essex, Colchester. Available at: https://www.iser.essex.ac.uk/files/bhps/quality-profiles/BHPS-QP-01-03-06-v2.pdf.

Lound, C. and Broad, P. (2013) *Initial Review of the FRS Weighting Scheme.* Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/321820/initial-review-family-resources-survey-weighting-scheme.pdf

Mayhew, L., Karlsson, M. and Rickayzen B. (2010) The Role of Private Finance in Paying for Long Term Care. *The Economic Journal*, **120**(548), F478–F504.

McVicar, D. (2008) Why have UK disability benefit rolls grown so much? *Journal of Economic Surveys* **22**, 114-139.

Melzer, D. Gardener, E. and Guralnik J. (2005) Mobility disability in the middle-aged: cross-sectional associations in the English Longitudinal Study of Ageing, *Age and Ageing*, **34** (6), 594-602.

Morciano, M. Hancock, R. and Pudney S. (2014) Disability costs and equivalence scales in the older population, *Review of Income and Wealth*, DOI: 10.1111/roiw.12108.

Muthén, L.K. and Muthén B.O. (2010) *Mplus User's Guide. Sixth Edition,* Muthén and Muthén, Los Angeles www.statmodel.com

OECD (2005) Long Term Care for Older People, Paris: OECD Publishing.

Oswald, A. J. and Powdthavee N., (2008) Does happiness adapt? A longitudinal study of disability with implications for economists and judges. *Journal of Public Economics*, **92**(5-6), 1061-1077.

Pickard, L, A. Comas-Herrera, J. Costa-Font, C. Gori, A. Di Maio, C. Patxot, A. Pozzi, H. and Wittenberg R. (2007) Modelling an Entitlement to Long-Term Care Services for Older People in Europe: Projections for Long-Term Care Expenditure to 2050. *Journal of European Social Policy* **17**, 33–48.

Pudney, S. Hancock, R. and Sutherland H. (2006) Simulating the reform of means-tested benefits with endogenous take-up and claim costs. *Oxford Bulleting of Economics and Statistics* **68** (2) 135-166

Pudney, S. E. (2010) *Disability benefits for older people: How does the UK Attendance Allowance system really work?* University of Essex: ISER Working Paper no. 2010-02.

Rässler, S. (2002) Statistical matching. A Frequentist Theory, Practical Applications, and Alternative Bayesian Approaches. Series: Lecture Notes in Statistics, Vol. 168, XVIII, Springer.

Rosenbaum, P.R. (2002) Observational Studies, Second Edition, Springer Series in Statistics, Springer.

Satorra, A. and Bentler P. M. (2001) A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika* **66**, 507–514.

Sen, A. K. (1985) *Commodities and Capabilities*. Amsterdam: North Holland.

Spector, W. D. and Fleishman, J. A. (1998), Combining activities of daily living with instrumental activities of daily living to measure functional disability, *Journal of Gerontology: Social Sciences*, **53B**(1), S46–S57.

Sturgis, P., Thomas, R., Purdon, S., Bridgwood, A. and Dodd T. (2001) *Comparative review and assessment of key health state measures of the general population*. Research report, London: Department of Health, UK.

Sutherland, S. (ed.) (1999) *With Respect to Old Age: Long-term Care – Rights and Responsibilities. A Report by the Royal Commission on Long Term Care*. London: The Stationery Office CM4192.

Taylor, M.F. (ed.), with Brice, J., Buck, N. and Prentice-Lane E. (2006) *British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices*. Institute for Social and Economic Research, University of Essex, Colchester. Available at http://www.iser.essex.ac.uk/ulsc/bhps/doc/pdf_versions/ volumes/bhpsvola.pdf.

Taylor, R., Conway, L., Calderwood, L., and Lessof C. (2003) Methodology. In *Health, wealth and lifestyles of the older population in England: The 2002 English Longitudinal Study of Ageing* (eds M. Marmot, J. Banks, R. Blundell, C. Lessof and J. Nazroo), London :The Institute for Fiscal Studies, pp. 357–374.

Taylor, R., Conway, L., Calderwood, L., Lessof, C., Cheshire, H., Cox, K. and Scholes S., (2007) *Technical report (ELSA wave 1): health, wealth and lifestyles of the older population in England*, London : National Centre for Social Research. Available at: http://www.ifs.org.uk/elsa/report03/w1_tech.pdf

Van de Ven W.P. and Van Der Gaag J. (1982) Health as an unobservable: A MIMIC-model of demand for health care. Journal of Health Economics, **1**, 157-83.

Verbrugge, L. M. (1980) Sex differences in complaints and diagnoses. *Journal of Behavioural Medicine,* **3**, 327-355.

Verbrugge, L. M. and Wingard D. L. (1987) Sex differentials in health and mortality. *Women's Health*, **12**,103-145.

Wallace R.B. and Herzog A. R. (1995) Overview of the Health Measures in the Health and Retirement Study. *The Journal of Human Resources*, Special Issue on the Health and Retirement Study: Data Quality and Early Results, **30**, S84-S107

Wanless, D. (2006) Securing Good Care for Older People: Taking a Long-Term View. London: King`s Fund. Available at: http://www.kings-fund.org.uk/sites/files/kf/field/field_publication_file/securing-good-care-for-older-people-wanless-2006.pdf

Wolfe B.L. and Behrman J.R. (1984) Determinants of Women's Health Status and Health-Care Utilization in a Developing Country: A Latent Variable Approach. *The Review of Economics and Statistics*, **66**, 696-703.

Zantomio, F. (2013) Older people's participation in extra-cost disability benefits. *Journal of Health Economics*, **32** (1), 320-330.

Appendix: Additional Tables

Table A1: Survey specific functional limitations indicators D

Data Source:		Not receiving AA un-weighted mean	Not receiving AA weighted mean	Receiving AA un-weighted mean	Receiving AA weighted mean	Non-recipient/ recipient difference (un-weighted)†
FRS: *Has difficulty with:*						
MOBILITY	mobility (moving about)	0.251	0.254	0.814	0.813	-0.563
LIFTING	lifting, carrying or moving objects	0.221	0.221	0.745	0.749	-0.524
DEXTERITY	manual dexterity using hands for everyday tasks	0.077	0.077	0.396	0.400	-0.319
CONTINENCE	with continence (bladder control)	0.055	0.056	0.237	0.235	-0.182
COMMUNICATION	communication (speech, hearing or eyesight)	0.039	0.040	0.204	0.200	-0.165
MEMORY	memory/concentration/learning/understanding	0.049	0.050	0.252	0.255	-0.203
KNOWING DANGER	recognising when in physical danger	0.005	0.005	0.068	0.069	-0.062
OTHER	other area of life	0.040	0.040	0.092	0.091	-0.053
PROXY	interviewed by proxy	0.059	0.059	0.121	0.131	-0.063
Observations		*6,093*		*651*		
ELSA: *Has difficulty with:*						
WALKING 100 YDS	walking 100 yards	0.117	0.121	0.572	0.582	-0.455
SITTING 2 HRS	sitting for about two hours	0.126	0.126	0.285	0.279	-0.158
CHAIR TRANSFERS	getting up from a chair after sitting for long periods	0.282	0.285	0.626	0.618	-0.344
STAIRS (several flights)	climbing several flights of stairs without resting	0.424	0.429	0.821	0.822	-0.397
STAIRS (1 flights)	climbing one flight of stairs without resting	0.161	0.167	0.650	0.653	-0.489
STOOPING	stooping, kneeling, or crouching	0.411	0.415	0.791	0.798	-0.381
REACHING	reaching or extending arms above shoulder level	0.103	0.105	0.344	0.339	-0.241
PULL/PUSHING	pulling or pushing large objects e.g. living room chair	0.183	0.189	0.675	0.686	-0.492
LIFTING	lifting/carrying weights over 10 lbs, e.g. heavy bag	0.281	0.288	0.797	0.806	-0.516
PICKING-UP COIN	picking up a 5p coin from a table	0.049	0.050	0.241	0.249	-0.192
DRESSING	ADL:dressing, including putting on shoes an	0.126	0.128	0.472	0.460	-0.346
WALKING	ADL:walking across a room	0.025	0.027	0.203	0.211	-0.178
BATHING	ADL:bathing or showering	0.128	0.132	0.566	0.568	-0.438
FEEDING	ADL:eating, such as cutting up your food	0.012	0.012	0.092	0.095	-0.08
BED TRANSFERS	ADL:getting in or out of bed	0.044	0.045	0.287	0.280	-0.243
USING TOILET	ADL:using the toilet, including getting up	0.029	0.030	0.179	0.179	-0.15
CONTINENCE	Problem with continence	0.157	0.158	0.336	0.338	-0.179
USING MAP	IADL:using a map to figure out how to get around	0.057	0.061	0.222	0.240	-0.165
PREP HOT MEAL	IADL:preparing a hot meal	0.029	0.031	0.282	0.291	-0.253
SHOPPING	IADL:shopping for groceries	0.083	0.088	0.504	0.515	-0.422
PHONING	IADL:making telephone calls	0.020	0.022	0.095	0.095	-0.075
MEDICATION	IADL:taking medications	0.010	0.011	0.084	0.086	-0.073
HOUSEWORK	IADL:doing work around the house or garden	0.159	0.163	0.650	0.660	-0.491
MANAGING MONEY	IADL: managing money, e.g. paying bills	0.023	0.025	0.154	0.162	-0.131
Observations		*4,773*		*369*		
BHPS: *Health hinders:*						
HOUSEWORK	doing the housework	0.089	0.095	0.573	0.557	-0.484
CLIMBING STAIRS	climbing the stairs	0.105	0.114	0.600	0.601	-0.495
DRESSING	getting dressed	0.036	0.038	0.173	0.185	-0.137
WALKING>10 mins	walking more than 10 mins	0.094	0.097	0.520	0.526	-0.426
How manages...(6-point scale)						
STAIRS	Stairs	1.856	1.914	3.920	3.830	-2.064
AROUND HOUSE	getting around house	1.350	1.367	2.613	2.551	-1.264
BED TRANSFERS	getting in/out bed	1.360	1.378	2.547	2.525	-1.187
CUTTING TOENAILS	cutting toenails	2.555	2.643	4.920	4.915	-2.365
BATHING	bathing/showering	1.572	1.626	3.280	3.286	-1.708
WALKING DOWN ROAD	walking down road	1.678	1.720	3.773	3.739	-2.095
Observations		*967*		*75*		

† All differences are significantly different from 0 at the 1% level.

Table A2: Estimates of the disability equation in weighted samples

	Coefficients			Tests and coefficient differences		
Covariates	FRS	ELSA	BHPS	FRS-ELSA	FRS-BHPS	ELSA-BHPS
Spline age 65-73	0.041†	0.033†	0.136†	0.007	-0.096†	-0.103†
	(0.013)	(0.013)	(0.038)	(0.018)	(0.041)	(0.04)
Spline from age 73+	0.093†	0.101†	0.119†	-0.008	-0.027	-0.019
	(0.008)	(0.008)	(0.02)	(0.011)	(0.022)	(0.021)
Post-compulsory education	-0.275†	-0.306†	-0.214	0.031	-0.061	-0.092
	(0.067)	(0.063)	(0.155)	(0.092)	(0.169)	(0.167)
Income spline to median	-0.138†	-0.055	-0.162§	-0.083	0.024	0.108
	(0.045)	(0.054)	(0.096)	(0.07)	(0.106)	(0.11)
Income spline from median	-0.354†	-0.276†	-0.599†	-0.078	0.246	0.323
	(0.086)	(0.075)	(0.218)	(0.114)	(0.235)	(0.231)
Outright owner	-0.369†	-0.482†	-0.14	0.113	-0.229	-0.342‡
	(0.065)	(0.066)	(0.16)	(0.093)	(0.173)	(0.173)
	3.004†	2.608†	3.376†	0.396	-0.372	-0.768
Variance (σ_v^2)	(0.281)	(0.238)	(0.823)	(1.075)	(0.428)	(0.896)
		Sample size			*Coefficient equality* ^{2}(6)	
	6,744	5,142	1,042	3.701	10.579	14.318‡

Statistical significance of the coefficient, t-test cross-sample coefficient difference and χ^2 statistic:
† p < 0.01; ‡ p < 0.05; § p < 0.1. Standard Errors in parenthesis.

Table A3: Estimates of the AA receipt equation in weighted samples

	Coefficients			Coefficient differences		
Covariates	FRS	ELSA	BHPS	FRS-ELSA	FRS-BHPS	ELSA-BHPS
Latent disability η	0.569†	0.467†	0.505†	0.101§	0.064	-0.037
	(0.042)	(0.035)	(0.092)	(0.055)	(0.101)	(0.099)
Female	0.144‡	0.238†	-0.047	-0.095	0.19	0.285
	(0.066)	(0.074)	(0.184)	(0.099)	(0.195)	(0.198)
Spline age 65-73	-0.043†	-0.036†	-0.088†	-0.007	0.045§	0.052‡
	(0.008)	(0.007)	(0.022)	(0.011)	(0.023)	(0.023)
Spline from age 73+	0.056†	0.041†	0.022	0.015	0.034‡	0.019
	(0.006)	(0.007)	(0.016)	(0.009)	(0.017)	(0.017)
Post- compulsory education	-0.148‡	-0.232†	-0.099	0.084	-0.049	-0.133
	(0.066)	(0.072)	(0.155)	(0.098)	(0.168)	(0.171)
(ln) income spline to median	-0.013	-0.071	0.002	0.057	-0.015	-0.072
	(0.049)	(0.053)	(0.097)	(0.072)	(0.109)	(0.111)
(ln) income spline from median	-0.432†	-0.405†	-0.375	-0.027	-0.056	-0.029
	(0.12)	(0.152)	(0.25)	(0.193)	(0.277)	(0.292)
Outright owner	-0.135‡	-0.019	-0.244	-0.116	0.109	0.225
	(0.063)	(0.074)	(0.171)	(0.097)	(0.183)	(0.187)
Married/cohabiting	-0.038	0.087	-0.105	-0.126	0.067	0.192
	(0.066)	(0.077)	(0.196)	(0.102)	(0.207)	(0.21)
$\chi^2(9)$ test of coefficient equality				13.015	17.027‡	13.457

Statistical significance of the coefficient, t-test cross-sample coefficient difference and χ^2 statistic:
† p < 0.01; ‡ p < 0.05; § p < 0.1. Standard Errors in parenthesis.

Table A4: Estimates of the disability equation in matched samples

Covariate	Coefficient Estimates (standard errors)			
FRS sample composition	***ELSA matched to FRS***		***BHPS matched to FRS***	
	FRS	ELSA	FRS	BHPS
Spline age 65-73	0.047	0.036	0.073	0.142
	(0.016)	(0.013)	(0.085)	(0.038)
Spline age 73+	0.090	0.098	0.077	0.119
	(0.010)	(0.008)	(0.034)	(0.020)
Post- compulsory education	-0.182	-0.231	-0.001	-0.090
	(0.082)	(0.066)	(0.264)	(0.161)
Income spline to median	-0.258	-0.113	-0.662	-0.925
	(0.097)	(0.094)	(0.667)	(0.381)
Income spline from median	-0.314	-0.391	-0.308	-0.469
	(0.122)	(0.089)	(0.323)	(0.251)
outright owner	-0.447	-0.491	-0.146	-0.226
	(0.082)	(0.068)	(0.167)	(0.167)
χ^2 (6) *for coefficient equality*	1.924		5.548	
Sample size	4,587		973	
ELSA sample composition	***FRS matched to ELSA***		***BHPS matched to ELSA***	
	FRS	ELSA	ELSA	BHPS
Spline age 65-73	0.033	0.037	0.061	0.072
	(0.016)	(0.013)	(0.031)	(0.035)
Spline age 73+	0.096	0.098	0.082‡	0.128‡
	(0.010)	(0.008)	(0.016)	(0.022)
Post- compulsory education	-0.205	-0.271	-0.043	-0.257
	(0.079)	(0.067)	(0.143)	(0.171)
Income spline to median	-0.125	-0.093	-0.284	-0.608
	(0.084)	(0.096)	(0.190)	(0.382)
Income spline from median	-0.340	-0.362	-0.245	-0.512
	(0.118)	(0.090)	(0.195)	(0.268)
outright owner	-0.437	-0.524	-0.442	-0.230
	(0.079)	(0.069)	(0.148)	(0.164)
χ^2 (6) *for coefficient equality*	1.548		6.241	
Sample size	4,596		850	
BHPS sample composition	***FRS matched to BHPS***		***ELSA matched to BHPS***	
	FRS	BHPS	ELSA	BHPS
Spline age 65-73	0.040‡	0.143‡	0.044‡	0.133‡
	(0.039)	(0.037)	(0.034)	(0.041)
Spline age 73+	0.089	0.116	0.089	0.112
	(0.021)	(0.020)	(0.019)	(0.021)
Post- compulsory education	-0.075	-0.053	0.112	-0.091
	(0.167)	(0.156)	(0.159)	(0.174)
Income spline to median	-0.444	-0.941	0.138	-0.296
	(0.425)	(0.367)	(0.296)	(0.266)
Income spline from median	-0.403	-0.423	-0.606	-0.551
	(0.252)	(0.249)	(0.275)	(0.301)
outright owner	-0.457	-0.209	-0.648	-0.318
	(0.182)	(0.161)	(0.172)	(0.183)
χ^2 (6) *for coefficient equality*	7.681		9.870	
Sample size	966		791	

Note: Significance of t-test cross-sample coefficient difference and χ^2 statistic: † $p < 0.01$; ‡ $p < 0.05$; § $p < 0.1$.

Table A5: Estimates of the AA receipt equation in matched samples

Covariate	Coefficient Estimates (standard errors)			
FRS sample composition	***ELSA matched to FRS***		***BHPS matched to FRS***	
	FRS	ELSA	FRS	BHPS
Latent disability η	0.550	0.498	0.622	0.517
	(0.047)	(0.038)	(0.117)	(0.094)
Female	0.031	0.179	-0.037	-0.128
	(0.083)	(0.080)	(0.181)	(0.186)
Spline age 65-73	-0.031	-0.025	-0.004	0.001
	(0.010)	(0.009)	(0.058)	(0.036)
Spline age 73+	0.062	0.050	0.023	0.025
	(0.008)	(0.007)	(0.021)	(0.016)
Post- compulsory education	-0.119	-0.209	-0.107	0.146
	(0.080)	(0.080)	(0.179)	(0.169)
Income spline to median	-0.125	-0.203	-0.349	-0.688
	(0.081)	(0.086)	(0.434)	(0.360)
Income spline from median	-0.398	-0.492	-0.644	-0.304
	(0.169)	(0.200)	(0.339)	(0.267)
outright owner	-0.113	0.010	-0.223	-0.297
	(0.077)	(0.078)	(0.173)	(0.171)
Married/Cohabiting	-0.010	0.079	0.110	-0.047
	(0.082)	(0.084)	(0.179)	(0.196)
χ^2 (9) *for coefficient equality*	6.447		3.000	
Sample size	4,587		973	
ELSA sample composition	***FRS matched to ELSA***		***BHPS matched to ELSA***	
	FRS	ELSA	ELSA	BHPS
Latent disability η	0.581	0.480	0.658	0.508
	(0.051)	(0.038)	(0.119)	(0.101)
Female	0.084	0.172	0.420	0.025
	(0.082)	(0.080)	(0.219)	(0.198)
Spline age 65-73	-0.028	-0.027	-0.037	-0.003
	(0.010)	(0.009)	(0.026)	(0.032)
Spline age 73+	0.057	0.050	0.057‡	0.021‡
	(0.008)	(0.007)	(0.019)	(0.017)
Post- compulsory education	-0.139	-0.207	-0.542	0.075
	(0.082)	(0.080)	(0.209)	(0.180)
Income spline to median	-0.154	-0.184	-0.241	-0.388
	(0.080)	(0.084)	(0.207)	(0.196)
Income spline from median	-0.415	-0.530	-0.525	-0.232
	(0.170)	(0.201)	(0.449)	(0.311)
outright owner	-0.089	0.027	-0.017	-0.251
	(0.078)	(0.078)	(0.192)	(0.178)
Married/Cohabiting	-0.066	0.084	0.023	-0.275
	(0.082)	(0.082)	(0.224)	(0.199)
χ^2 (9) *for coefficient equality*	7.590		11.522	
Sample size	4,596		850	

BHPS sample composition	*FRS matched to BHPS*		*ELSA matched to BHPS*	
	FRS	BHPS	ELSA	BHPS
Latent disability η	0.519	0.530	0.566	0.510
	(0.098)	(0.096)	(0.100)	(0.103)
Female	-0.115	-0.131	0.059	-0.128
	(0.171)	(0.184)	(0.202)	(0.184)
Spline age 65-73	-0.005	0.001	-0.038	-0.047
	(0.032)	(0.035)	(0.023)	(0.030)
Spline age 73+	0.048	0.026	0.057	0.032
	(0.017)	(0.016)	(0.017)	(0.017)
Post- compulsory education	-0.076	0.147	-0.388	0.050
	(0.171)	(0.171)	(0.210)	(0.175)
Income spline to median	-0.223	-0.692	-0.265	-0.381
	(0.335)	(0.360)	(0.207)	(0.206)
Income spline from median	-0.524	-0.334	0.131	-0.318
	(0.374)	(0.27)	(0.383)	(0.308)
outright owner	-0.259	-0.302	0.011	-0.289
	(0.176)	(0.171)	(0.202)	(0.183)
Married/Cohabiting	-0.021	-0.031	-0.095	-0.103
	(0.200)	(0.195)	(0.207)	(0.198)
χ^2 (9) *for coefficient equality*	3.445		6.619	
Sample size	966		791	

Note: Significance of t-test cross-sample coefficient difference and χ^2 statistic: † $p < 0.01$; ‡ $p < 0.05$; § $p < 0.1$.

Table A6: Estimates of the latent disability equation for the FRS and ELSA 2-factor models

Covariates	1			2		
	FRS	ELSA	Tests and coefficient differences	FRS	ELSA	Tests and coefficient differences
Spline age 65-73	0.033†	0.035†	-0.002	0.025	-0.015	0.040§
	(0.003)	(0.011)	(0.011)	(0.016)	(0.013)	(0.021)
Spline from age 73+	0.064†	0.095†	-0.031†	0.079†	0.071†	0.008
	(0.005)	(0.007)	(0.009)	(0.008)	(0.009)	(0.012)
Post-compulsory education	-0.237†	-0.276†	0.039	-0.142§	-0.241†	0.100
	(0.051)	(0.058)	(0.077)	(0.075)	(0.069)	(0.102)
Income spline to median	-0.103†	-0.039	-0.063	-0.175†	-0.119‡	-0.056
	(0.037)	(0.051)	(0.063)	(0.038)	(0.047)	(0.061)
Income spline from median	-0.293†	-0.305†	0.013	-0.086	-0.170§	0.084
	(0.071)	(0.070)	(0.100)	(0.102)	(0.090)	(0.136)
Outright owner	-0.334†	-0.484†	0.150§	-0.120§	-0.135‡	0.015
	(0.053)	(0.062)	(0.081)	(0.072)	(0.061)	(0.095)
^{2}(6) coefficient equality			19.616†			7.423
Sample size	6744			5142		

Statistical significance of the coefficient, t-test cross-sample coefficient difference and χ^2 statistic:
† $p < 0.01$; ‡ $p < 0.05$; § $p < 0.1$. Standard Errors in parenthesis.

Table A7: Estimates of the AA receipt equation for the FRS and ELSA 2-factor models

Covariates	FRS		ELSA		tests and coefficient differences	
Latent disability η_1	0.508†	(0.039)	0.419†	(0.045)	0.089	(0.060)
Latent disability η_2	0.295†	(0.046)	0.164§	(0.089)	0.131	(0.100)
Female	-0.043†	(0.006)	-0.032†	(0.007)	-0.012	(0.010)
Spline age 65-73	0.055†	(0.006)	0.042†	(0.007)	0.013	(0.009)
Spline from age 73+	-0.166‡	(0.065)	-0.222†	(0.072)	0.056	(0.097)
Post- compulsory education	-0.001	(0.048)	-0.078	(0.050)	0.077	(0.069)
(ln) income spline to median e	-0.406†	(0.120)	-0.421†	(0.153)	0.015	(0.195)
(ln) income spline from median	-0.149‡	(0.063)	-0.015	(0.072)	-0.135	(0.096)
Outright owner	-0.079	(0.065)	0.084	(0.077)	-0.163	(0.101)
Married/cohabiting	0.183‡	(0.072)	0.271†	(0.075)	-0.088	(0.104)
^{2}(10) coefficient equality					22.477‡	
Sample size	6744		5142			

Statistical significance of the coefficient, t-test cross-sample coefficient difference and χ^2 statistic:
† $p < 0.01$; ‡ $p < 0.05$; § $p < 0.1$. Standard Errors in parenthesis.

Online Appendix: Further Tables and Identification proof

Table O1: Sample means of SES and AA receipt in FRS, ELSA and BHPS

	FRS		ELSA		BHPS	
	mean	sd	mean	sd	mean	sd
			Unweighted			
Female	0.559	*0.497*	0.557	*0.497*	0.560	*0.497*
Age†	74		73		74	
Post-compulsory education	0.505	*0.500*	0.539	*0.499*	0.513	*0.500*
Ln pre-benefit equivalised income‡	6.454	*0.806*	6.412	*0.751*	6.551	*0.732*
Outright owner	0.664	*0.472*	0.690	*0.463*	0.701	*0.458*
Married/cohabiting	0.579	*0.494*	0.565	*0.496*	0.553	*0.497*
Receives AA	0.097	*0.295*	0.072	*0.258*	0.072	*0.259*
			Weighted			
Female	0.555	0.497	0.571	0.495	0.561	0.497
Age†	74		74		74	
Post-compulsory education	0.513	0.500	0.522	0.500	0.495	0.500
Ln pre-benefit equivalised income‡	6.463	0.826	6.391	0.754	6.521	0.746
Outright owner	0.677	0.468	0.682	0.466	0.672	0.470
Married/cohabiting	0.573	0.495	0.548	0.498	0.538	0.499
Receives AA	0.094	0.292	0.077	0.266	0.079	0.270
Observations	6,746		5,142		1,042	

Notes: †To protect confidentiality, FRS and ELSA release data with a top-coding at the age of 80 and 90, respectively. Therefore, we report median rather than mean values. ‡ Household income excludes disability and means tested benefits and it has been equivalised using the modified OECD equivalence scale.

Table O2: Sample means of SES and AA receipt in matched samples

	FRS		ELSA		FRS		BHPS	
	mean	sd	mean	sd	mean	sd	mean	sd
FRS sample composition:	*ELSA matched to FRS*				*BHPS matched to FRS*			
	FRS		ELSA		FRS		BHPS	
Female	0.561	*0.496*	0.561	*0.496*	0.566	*0.496*	0.566	*0.496*
Age†	73		73		74		74	
Post-compulsory schooling	0.530	*0.499*	0.530	*0.499*	0.506	*0.500*	0.506	*0.500*
ln pre-benefit equivalised income‡	6.457	*0.582*	6.456	*0.582*	6.576	*0.503*	6.600	*0.500*
Accommodation own it outright	0.690	*0.462*	0.690	*0.462*	0.716	*0.451*	0.716	*0.451*
Married/cohabiting	0.572	*0.495*	0.572	*0.495*	0.565	*0.496*	0.565	*0.496*
Receives AA	0.088	*0.283*	0.071	*0.257*	0.094	*0.291*	0.072	*0.259*
Observations	4,587				973			
ELSA sample composition:	*FRS matched to ELSA*				*BHPS matched to ELSA*			
	FRS		ELSA		ELSA		BHPS	
Female	0.562	*0.496*	0.562	*0.496*	0.575	*0.495*	0.575	*0.495*
Age†	73		73		74		74	
Post-compulsory schooling	0.531	*0.499*	0.531	*0.499*	0.504	*0.500*	0.504	*0.500*
ln pre-benefit equivalised income‡	6.458	*0.578*	6.455	*0.582*	6.563	*0.513*	6.533	*0.527*
accommodation own it outright	0.690	*0.463*	0.690	*0.463*	0.720	*0.449*	0.720	*0.449*
Married/cohabiting	0.574	*0.495*	0.574	*0.495*	0.552	*0.498*	0.552	*0.498*
Receives AA	0.089	*0.284*	0.070	*0.255*	0.072	*0.258*	0.066	*0.248*
Observations	4,596				850			
BHPS sample composition:	*FRS matched to BHPS*				*ELSA matched to BHPS*			
	FRS		BHPS		ELSA		BHPS	
Female	0.565	*0.496*	0.565	*0.496*	0.564	*0.496*	0.564	*0.496*
Age†	74		74		74		74	
Post-compulsory schooling	0.505	*0.500*	0.505	*0.500*	0.497	*0.500*	0.497	*0.500*
ln pre-benefit equivalised income‡	6.575	*0.499*	6.599	*0.496*	6.488	*0.496*	6.513	*0.500*
accommodation own it outright	0.716	*0.451*	0.716	*0.451*	0.718	*0.450*	0.718	*0.450*
Married/cohabiting	0.566	*0.496*	0.566	*0.496*	0.550	*0.498*	0.550	*0.498*
Receives AA	0.085	*0.279*	0.072	*0.259*	0.068	*0.252*	0.078	*0.269*
Observations	966				791			

Notes: Based on unweighted selected samples. † To protect confidentiality, FRS and ELSA release data with a top-coding at the age of 80 and 90, respectively. Therefore, we report median rather than mean values. ‡ Household income excludes disability and means tested benefits and it has been equivalised using the modified OECD equivalence scale.

Table O3: Factor loadings for the FRS and ELSA 2-factor models and squared correlations of disability indicators with latent indices ($\boldsymbol{\eta_q}$)

Functional limitation indicator	Male		Female	
	Factor 1 (η_1)	Factor 2 (η_2)	Factor 1 (η_1)	Factor 2 (η_2)
FRS $cov(\eta_1, \eta_2)$	1.172		0.854	
MOBILITY	1		1	
LIFTING	1.586†		2.226†	
DEXTERITY	0.768†		0.736†	
CONTINENCE	0.315†	0.235†	0.363†	0.275†
COMMUNIC		1		1
MEMORY		0.837†		0.987†
DANGER		1.005†		1.078†
OTHER	0.009	0.144‡	-0.064	0.208†
PROXY		0.204†		0.270†
ELSA $cov(\eta_1, \eta_2)$	1.058		0.890	
WALK100	1		1	
SITTING	0.394†		0.409†	
CHAIR	0.593†		0.545†	
CLIMBSEV	0.736†		0.689†	
CLIMB1	1.014†		0.918†	
STOOP	0.657†		0.669†	
ARMS	0.511†		0.511†	
PULL/PUSH	1.025†		0.921†	
LIFTING	0.954†		0.919†	
COIN	0.383†		0.44†	
DRESSING	0.673†		0.665†	
WALKING	1.082†		0.980†	
BATH	0.879†		0.736†	
EATING	0.586†		0.431†	
BED	0.897†		0.705†	
TOILET	0.751†		0.592†	
CONTINENCE	0.196†	0.235‡	0.275†	-0.047
MAP		1.052†		1.031†
MEAL				
SHOPPING	0.999†		1.129†	
PHONE		1		1
MEDICATION		1.231†		1.319†
HOUSEWORK	1.137†		0.938†	
MONEY		1.25†		1.731†

Statistical significance of the factor loadings: † $p < 0.01$; ‡ $p < 0.05$; § $p < 0.1$.

Table O4: Factor loadings for the FRS and ELSA 1-factor models with alternative factor loading constraints

FRS					ELSA				
Disability Indicator	Factor loading (St. err.)				Disability Indicator	Factor loading (St. err.)			
	Men		Women			Men		Women	
MOBILITY	0.849†	(0.072)	0.962†	(0.077)	WALKING 100 YDS	1.118†	(0.079)	1.077†	(0.039)
LIFTING	1	-	1	-	SITTING 2 HRS	0.422†	(0.034)	0.436†	(0.030)
DEXTERITY	0.663†	(0.058)	0.579†	(0.040)	CHAIR TRANSFERS	0.635†	(0.042)	0.582†	(0.035)
CONTINENCE	0.360†	(0.035)	0.392†	(0.033)	STAIRS (several flights)	0.792†	(0.050)	0.735†	(0.042)
COMMUNIC	0.351†	(0.039)	0.333†	(0.035)	STAIRS (1 flight)	1.084†	(0.069)	0.984†	(0.058)
MEMORY	0.382†	(0.04)	0.380†	(0.035)	STOOPING	0.701†	(0.044)	0.715†	(0.040)
DANGER	0.461†	(0.086)	0.388†	(0.050)	REACHING	0.550†	(0.044)	0.547†	(0.037)
OTHER	0.089†	(0.025)	0.055‡	(0.022)	PULL/PUSHING	1.100†	(0.071)	0.987†	(0.050)
PROXY	0.105†	(0.027)	0.110†	(0.022)	LIFTING	1	-	1	-
					PICKING-UP COIN	0.415†	(0.051)	0.474†	(0.039)
					DRESSING	0.723†	(0.051)	0.711†	(0.046)
					WALK ACROSS ROOM	1.154†	(0.151)	1.048†	(0.099)
					BATHING	0.944†	(0.073)	0.790†	(0.050)
					FEEDING	0.652†	(0.093)	0.468†	(0.060)
					BED TRANSFERS	0.962†	(0.093)	0.751†	(0.058)
					USING TOILET	0.808†	(0.097)	0.631†	(0.054)
					CONTINENCE	0.327†	(0.032)	0.275†	(0.023)
					USING A MAP	0.445†	(0.052)	0.375†	(0.031)
					PREP. HOT MEAL	0.883†	(0.109)	0.886†	(0.081)
					SHOPPING	1.115†	(0.091)	1.241†	(0.086)
					PHONING	0.392†	(0.049)	0.357†	(0.049)
					MEDICATION	0.523†	(0.077)	0.524†	(0.081)
					HOUSEWORK	1.239†	(0.092)	1.014†	(0.063)
					MANAGING MONEY	0.496†	(0.061)	0.524†	(0.052)
Sample size	6,744					5,142			

Statistical significance of the factor loadings: † $p < 0.01$; ‡ $p < 0.05$; § $p < 0.1$.

Table O5: Estimates of the latent disability equation for the FRS and ELSA 1-factor models with alternative factor loading constraints

Covariates	Coefficients and Standard Errors: FRS	Coefficients and Standard Errors: ELSA	Tests and coefficient differences	
Spline age 65-73	0.042† (0.014)	0.032† (0.011)	0.010	(0.018)
Spline from age 73+	0.100† (0.009)	0.090† (0.007)	0.010	(0.011)
Post-compulsory education	-0.307† (0.074)	-0.255† (0.055)	-0.052	(0.092)
Income spline to median	-0.180† (0.052)	-0.042 (0.048)	-0.137§	(0.070)
Income spline from median	-0.369† (0.094)	-0.284† (0.066)	-0.085	(0.115)
Outright owner	-0.416† (0.071)	-0.444† (0.057)	0.028	(0.092)
	Sample size		*Coefficient equality* χ^2 (6)	
	6,744	5,142	7.573	

Statistical significance of the coefficient, t-test cross-sample coefficient difference and χ^2 statistic:
† $p < 0.01$; ‡ $p < 0.05$; § $p < 0.1$. Standard Errors in parenthesis.

Table O6: Estimates of the AA receipt equation for the FRS and ELSA 1-factor models with alternative factor loading constraints

	Coefficients and Standard Errors		Tests of coefficient equality	
Covariates	FRS	ELSA	FRS-ELSA	
Latent disability η	0.516† (0.041)	0.522† (0.038)	-0.006	(0.056)
Female	0.118§ (0.065)	0.252† (0.073)	-0.134	(0.098)
Spline age 65-73	-0.040† (0.008)	-0.036† (0.007)	-0.004	(0.011)
Spline from age 73+	0.058† (0.006)	0.046† (0.007)	0.012	(0.009)
Post- compulsory education	-0.161‡ (0.065)	-0.238† (0.071)	0.077	(0.096)
(ln) income spline to median	-0.007 (0.048)	-0.092§ (0.049)	0.085	(0.069)
(ln) income spline from median	-0.390† (0.120)	-0.422† (0.154)	0.032	(0.195)
Outright owner	-0.138‡ (0.062)	-0.006 (0.071)	-0.132	(0.095)
Married/cohabiting	-0.077 (0.064)	0.087 (0.076)	-0.164	(0.100)
	Sample size		*^{2}(9) test of coefficient equality*	
	6,744	5,142	10.841	

Statistical significance of the coefficient, t-test cross-sample coefficient difference and χ^2 statistic:
† $p < 0.01$; ‡ $p < 0.05$; § $p < 0.1$. Standard Errors in parenthesis.

Table O7: Factor loadings for the FRS 1-factor model excluding proxy cases from the FRS sample (and the proxy indicator from the measurement model)

	FRS			
	Factor loading (St. err.)			
Disability Indicator	Men		Women	
MOBILITY	1	-	1	-
LIFTING	1.039†	(0.103)	1.203†	(0.123)
DEXTERITY	0.683†	(0.065)	0.602†	(0.049)
CONTINENCE	0.343†	(0.036)	0.426†	(0.037)
COMMUNIC	0.338†	(0.041)	0.317†	(0.036)
MEMORY	0.356†	(0.039)	0.382†	(0.036)
DANGER	0.355†	(0.091)	0.408†	(0.063)
OTHER	0.101†	(0.029)	0.068†	(0.026)
Sample size	6,308			

Statistical significance of the factor loadings: † $p < 0.01$; ‡ $p < 0.05$; § $p < 0.1$.

Table O8: Estimates of the latent disability equations obtained by dropping proxy cases from the FRS sample (and the proxy indicator from the measurement model)

	Coefficients and Standard Errors			Tests and coefficient differences		
Covariates	FRS	ELSA§§	BHPS§§	FRS-ELSA	FRS-BHPS	ELSA-BHPS§§
Spline age 65-73	0.039†	0.035†	0.127†	0.003	-0.089‡	-0.092†
	(0.014)	(0.012)	(0.036)	(0.018)	(0.038)	(0.038)
Spline from age 73+	0.084†	0.099†	0.128†	-0.015	-0.044‡	-0.029
	(0.008)	(0.008)	(0.020)	(0.011)	(0.022)	(0.022)
Post-compulsory education	-0.301†	-0.280†	-0.182	-0.021	-0.119	-0.097
	(0.068)	(0.061)	(0.149)	(0.091)	(0.164)	(0.161)
Income spline to median	-0.114‡	-0.046	-0.172§	-0.068	0.057	0.125
	(0.052)	(0.052)	(0.104)	(0.074)	(0.116)	(0.116)
Income spline from median	-0.317†	-0.310†	-0.558†	-0.007	0.241	0.248
	(0.088)	(0.072)	(0.206)	(0.114)	(0.224)	(0.218)
Outright owner	-0.389†	-0.487†	-0.185	0.098	-0.204	-0.302§
	(0.067)	(0.064)	(0.151)	(0.092)	(0.165)	(0.163)
	Sample size			*Coefficient equality* χ^2 (6)		
	6,308	5,142	1,042	3.411	13.27‡	14.139‡

Statistical significance of the coefficient, t-test cross-sample coefficient difference and χ^2 statistic: † $p < 0.01$; ‡ $p < 0.05$; § $p < 0.1$. Standard Errors in parenthesis. §§ Estimates are the same reported in Table 3.

Table O9: Estimates of the AA receipt equations obtained by dropping PROXY cases from the FRS sample (and the proxy indicator from the measurement model)

	Coefficients and Standard Errors			Tests of coefficient equality		
Covariates	FRS	ELSA§§	BHPS§§	FRS-ELSA	FRS-BHPS	ELSA-BHPS§§
Latent disability η	0.573†	0.477†	0.538†	0.096§	0.036	-0.06
	(0.043)	(0.035)	(0.095)	(0.056)	(0.104)	(0.101)
Female	0.156‡	0.251†	-0.068	-0.095	0.224	0.319§
	(0.069)	(0.073)	(0.172)	(0.101)	(0.185)	(0.187)
Spline age 65-73	-0.041†	-0.036†	-0.084†	-0.004	0.043§	0.048‡
	(0.009)	(0.007)	(0.021)	(0.011)	(0.023)	(0.022)
Spline from age 73+	0.059†	0.046†	0.028§	0.014	0.031§	0.017
	(0.006)	(0.007)	(0.015)	(0.009)	(0.016)	(0.016)
Post- compulsory education	-0.153‡	-0.238†	-0.070	0.085	-0.083	-0.167
	(0.069)	(0.071)	(0.155)	(0.099)	(0.170)	(0.171)
(ln) income spline to median	-0.044	-0.092§	-0.041	0.048	-0.002	-0.050
	(0.063)	(0.049)	(0.090)	(0.079)	(0.109)	(0.102)
(ln) income spline from median	-0.493†	-0.422†	-0.411§	-0.071	-0.082	-0.011
	(0.136)	(0.154)	(0.247)	(0.205)	(0.282)	(0.291)
Outright owner	-0.137‡	-0.006	-0.265	-0.131	0.128	0.259
	(0.065)	(0.071)	(0.164)	(0.097)	(0.176)	(0.178)
Married/cohabiting	-0.058	0.087	-0.171	-0.145	0.112	0.257
	(0.068)	(0.076)	(0.182)	(0.102)	(0.195)	(0.198)
	Sample size			*Coefficient equality* χ^2 (9)		
	6,308	5,142	1,042	13.287	14.957§	14.844§

Statistical significance of the coefficient, t-test cross-sample coefficient difference and χ^2 statistic: † $p < 0.01$; ‡ $p < 0.05$; § $p < 0.1$. Standard Errors in parenthesis. §§ Estimates are the same reported in Table 4.

Identification

After using equation (3) to solve out the latent disability variables η_{iq} from the model, the structure can be written in matrix form as:

$$\tilde{\mathbf{D}} = \mathbf{\Lambda\Theta z} + \mathbf{\Lambda\upsilon} + \boldsymbol{\varepsilon} \qquad \text{(A1)}$$

$$\tilde{R} = (\boldsymbol{\beta} + \boldsymbol{\gamma\Theta})\mathbf{z} + \boldsymbol{\gamma\upsilon} + u \qquad \text{(A2)}$$

where $\boldsymbol{\Lambda}$, $\boldsymbol{\Theta}$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are respectively $K \times Q$, $Q \times p$, $1 \times p$ and $1 \times Q$ dimensional coefficient matrices and we have omitted the individual i suffix from the covariates $\mathbf{z}$, the latent variables $\tilde{\mathbf{D}}$, and $\tilde{R}$ underlying the observed ordinal variables $\mathbf{D}$ and R, and the unobservable random terms $\boldsymbol{\upsilon}$, $\boldsymbol{\varepsilon}$ and u. Equations (A1)-(A2) together comprise a system of correlated reduced form (ordered) probit equations, from which we can identify the following coefficient matrices and residual covariances:

$$\mathbf{B}_1 = \mathbf{\Lambda\Theta} \qquad \text{(A3)}$$

$$\mathbf{B}_2 = \boldsymbol{\beta} + \boldsymbol{\gamma\Theta} \qquad \text{(A4)}$$

$$\mathbf{C}_{11} = \mathbf{\Lambda\Omega\Lambda}' + \mathbf{\Sigma} \qquad \text{(A5)}$$

$$\mathbf{C}_{22} = \boldsymbol{\gamma\Omega\gamma}' + \boldsymbol{\gamma\delta} + \sigma^2 \qquad \text{(A6)}$$

$$\mathbf{C}_{12} = \mathbf{\Lambda\Omega}\boldsymbol{\gamma}' + \mathbf{\Lambda}\boldsymbol{\delta} \qquad \text{(A7)}$$

where $\boldsymbol{\Omega}$ is the covariance matrix of $\boldsymbol{\upsilon}$, $\boldsymbol{\Sigma}$ is the diagonal covariance matrix of $\boldsymbol{\varepsilon}$, $\boldsymbol{\delta}$ is the vector of covariances between $\boldsymbol{\upsilon}$ and u, and σ^2 is the variance of u.

Some normalisations are necessary, because the observed variables $\boldsymbol{D}$ and R do not reveal the scale of $\tilde{\mathbf{D}}$ and $\tilde{R}$ and because the latent $\boldsymbol{\eta}$ can be replaced by arbitrary linear combinations with the loadings $\boldsymbol{\Theta}$ and $\boldsymbol{\gamma}$ transformed accordingly. Without loss of generality, we resolve these indeterminacies by setting C_{22} and the diagonal elements of C_{11} to unity and by imposing the restrictions:

$$\boldsymbol{\Lambda} = \begin{pmatrix} \mathbf{I} \\ \boldsymbol{\Lambda}_2 \end{pmatrix} \tag{A8}$$

Given these normalisations, the first Q rows of $\boldsymbol{B}_1$ identify $\boldsymbol{\Theta}$. Provided the rank of $\boldsymbol{\Theta}$ is Q, $\boldsymbol{\Lambda}_2$ can then be found by solving the last K_s-Q equations in (A3). This rank condition implies that the Q latent factors in the measurement equations (1) cannot be replaced by a smaller number of linear combinations of the factors.

Now consider identification of $\boldsymbol{\Omega}$. Write the vector of Q diagonal elements of $\boldsymbol{\Omega}$ as $\boldsymbol{\omega}_{\mathrm{d}}$ and the vector of $(Q-1)/2$ sub-diagonal elements as $\boldsymbol{\omega}_{\mathrm{s}}$. We can construct an identity: vec($\boldsymbol{\Omega}$) = $\mathrm{S_d}\ \boldsymbol{\omega}_{\mathrm{d}} + \mathrm{S_s}\ \boldsymbol{\omega}_{\mathrm{s}}$ where $\mathrm{S} = (\mathrm{S_d}\ \mathrm{S_s})$ is a $Q^2 \times Q(Q+1)/2$ permutation matrix containing 1s and 0s and vec(.) is the operation of stacking the rows of a matrix into a column vector. Let $\mathbf{C}_{11}^{1,1}$ be the leading $Q \times Q$ block of $\mathbf{C}_{11}$ and note that $\boldsymbol{\Sigma}$ is diagonal so that $\mathbf{S}_s'\mathrm{vec}\left(\mathbf{C}_{11}^{1,1}\right) = \boldsymbol{\omega}_s$. This determines the off-diagonal elements of $\boldsymbol{\omega}$. Now let $\mathbf{C}_{11}^{1,2}$ be the submatrix of $\mathbf{C}_{11}$ containing elements from the first Q rows and last K_s-Q columns: then $\mathbf{C}_{11}^{1,2} = \boldsymbol{\Omega}\boldsymbol{\Lambda}_2{}'$ and, if c_{qj} is the typical element of $\mathbf{C}_{11}^{1,2}$, each of the ω_{qq} can be deduced as

$\omega_{qq} = \left(c_{qj} - \sum_{r \neq q} \omega_{qr} \lambda^s_{jr} \right) / \lambda^s_{jq}$, provided there exists at least one non-zero element in the qth column of $\boldsymbol{\Lambda}_2$, for each $q = 1 \ldots Q$. With $\boldsymbol{\Omega}$ determined, $\boldsymbol{\Sigma}$ is immediately given by (A5).

Without further restrictions, this is as far as we can go. Once $\boldsymbol{\Theta}$, $\boldsymbol{\Lambda}$, $\boldsymbol{\Omega}$ and $\boldsymbol{\Sigma}$ are known, this still leaves $p + 2Q + 1$ parameters , $\boldsymbol{\gamma}$, $\boldsymbol{\delta}$ and σ^2 to be determined by the $p + Q + 1$ equations in (A4), (A6) and (A7). At least Q further restrictions are necessary. Natural possibilities are $\boldsymbol{\delta} = cov(\boldsymbol{\upsilon}, u) = \mathbf{0}$ or exclusion restrictions on the vector $\boldsymbol{\beta}$. The latter requires the existence of covariates that can be assumed a priori to influence disability status (relevance) but have no causal role in determining benefit receipt (validity).

Chapter 3:

Disability costs and equivalence scales in the older population in Great Britain *

Abstract: We use a standard of living (SoL) approach to estimate older people's disability costs, using data on 8,000 individuals from the UK Family Resources Survey. We extend previous research in two ways. First, by allowing for a more flexible relationship between SoL and income, the structure of the estimated disability cost and equivalence scale is not dictated by a restrictive functional form assumption. Second, we allow for the latent nature of disability and SoL, addressing measurement error in the disability and SoL indicators in surveys. We find that disability costs are strongly related to severity of disability, and vary with income in absolute and proportionate terms. Older people above the median disability level require an extra £99 per week (2007 prices) on average to reach the standard of living of an otherwise similar person at the median. Costs faced by older people in the highest decile of disability average £180.

Keywords: costs of disability, disability indices, standard of living, equivalence scale, structural equation modelling.

JEL codes: C81, D10, I10.

---

* This chapter has been published in the *Review of Income and Wealth* (Volume 61, Issue 3, pp. 494–514, September 2015; DOI: 10.1111/roiw.12108) and it is a joint work with Ruth Hancock and Stephen Pudney. An earlier version is available as ISER working paper 2012-09, Colchester: Institute for Economic and Social Research, University of Essex (https://www.iser.essex.ac.uk/publications/working-papers/iser/2012-09). M. Morciano originated the study, conceptualized ideas, synthesized analyses, and interpreted findings. R. Hancock derived the dataset from the Family Resources Survey (FRS). R. Hancock and S. Pudney supervised all aspects of its implementation. All authors contributed to the writing of the article and reviewing drafts. Earlier versions of this study were presented at International Conference on Evidence-Based Policy in Long-Term Care, January 2012, London School of Economics, London; ESRC Research Centre on Micro-social Change (MISOC) research Workshop, 9 January 2012, Colchester; 4th Conference of the Society for the Study of Economic Inequality (ECINEQ), 18-20 July 2011, Catania (Italy). This work was supported by the Nuffield Foundation, by the Australian Research Council and by the Economic and Social Research Council through the Research Centre on Micro-social Change (MiSoC). Data from the Family Resources Survey (FRS) are made available by the UK Department of Work and Pensions through the UK Data Archive. Material from the FRS is Crown Copyright and is used by permission. Neither the collectors of the data nor the UKDA bear any responsibility for the analyses or interpretations presented here.

1 Introduction

Disabled people experience significant additional costs as a consequence of their disability. This is recognized in social security systems through the provision of benefits designed to compensate for disability-related consumption costs. There is no consensus on the scale of these costs (Stapleton *et al.*, 2008) and thus it is hard to assess how far social security systems compensate for them in practice. In the UK, older people with disabilities may be entitled to one of two social security benefits which are intended to help with the extra costs of disability: *Attendance Allowance* (AA) and *Disability Living Allowance* (DLA). AA can be claimed only by people aged 65 and over; DLA must be claimed before reaching age 65, but if awarded, can continue past age 65.[36] AA is paid at one of two rates depending on level of disability or care needs. DLA has a care component and a mobility component. The care component is payable at one of three levels corresponding to different degrees of care need; the mobility component is paid at one of two rates according to mobility needs.[37] About a quarter of people aged 65 and over receive AA or DLA (Hancock and Pudney, 2013). The benefits are not means tested although they can trigger additional entitlements to means-tested benefits

[36] From April 2013, DLA will start to be replaced by Person Independence Payment which will differ from DLA in certain details (Welfare Reform Act 2012)

[37] In 2007, the year to which our data relate, the two rates of AA were £64.50 or £43.15. In 2007 the levels of DLA were such that weekly payments ranged from £17.75 to £109.50.

through a Severe Disability Premium.[38] People with care needs may also be entitled to publicly-funded and largely means-tested social care in their own homes or in care homes. Such care is received by only 6% of the older population (Wittenberg *et al.*, 2011).There is continuing international debate on how best to fund the care needs of growing numbers of older people (Da Roit and Le Bihan, 2010; Gleckman, 2010; Swartz *et al.*, 2012). The role of cash disability benefits in the overall system of public support for care needs is an important part of this debate. It is therefore important to have methods to derive evidence on the extent to which the levels of cash disability benefits compensate for the extra costs that different degrees of disability bring. Moreover, when carrying out analysis of the distributional impact of tax and social security benefit reforms, it is crucially important to make some allowance for these additional living costs. If disability benefits are included in income, failure to do so would give a misleadingly favourable view of the position of disabled people in the income distribution (Hancock and Pudney, 2013).

At least five different methods have been used to estimate and adjust for the costs of disability. One is to exploit the existing benefit system and assume that the political process has resulted in an acceptable evaluation of disability costs. This implies use of an income measure for distributional analysis which excludes

[38] Worth up to £48.45 in 2007 for an older disabled person receiving a means-tested benefit.

any receipt of disability benefit (see Hancock and Pudney, 2013; Hancock *et al.*, 2013), on the assumption that income from disability benefit is exactly offset by the extra costs of disability. However, in practice such payments follow simple rules not well tailored to each individual's specific configuration of impairments and they are not necessarily intended to meet the full costs of disability. There may also be imperfections in the eligibility judgements made by programme administrators and non take-up by potential claimants. Consequently, this approach may give a poor approximation to disability costs, with underestimation in many cases, leading to bias in distributional analysis. Clearly it cannot be used to assess the adequacy of existing disability benefit levels.

A second, judgement-based, approach attempts to estimate the disability costs by asking a panel of 'experts', or disabled people themselves, to identify disability-related costs: see Martin and White, 1988, Thompson *et al.*, 1990, Smith *et al.*, 2004 for examples of this approach. The difficulty here is that the appropriate costs may depend not only on the nature of the impairments suffered by the individual, but also other characteristics that vary across households, and it is not feasible to use expert judgement at the level of individual respondents to large-scale surveys. Disabled people themselves may also find it difficult to envisage and evaluate the counterfactual situation in which their disability is removed but all else remains constant.

A third 'objective' revealed preference approach constructs an equivalence scale by using the consumption pattern (typically the household's food budget share) as an indicator of living standards in a comparison of a sample of disabled people with matched individuals who are unaffected by disability. This has been done extensively in the context of adjustment for household size and structure, but less often for disability (although see Jones and O'Donnell, 1995 for a UK example). The main difficulty with this revealed preference method is the need for strong assumptions to overcome inherent identification problems (Pollack and Wales, 1979; Muellbauer, 1979; Coulter *et al.*, 1992; Banks *et al.*, 1997; Deaton and Paxson, 1998).

A fourth alternative is to use a 'subjective' equivalence approach, based on individuals' reported satisfaction with their well-being. Two main types of subjective information have been used: evaluations of standard of living using an arbitrary numerical scale; or judgements on the level of income believed necessary to reach a specified standard of living (see Stewart, 2009). For the subjective approach, there are concerns about the quality of subjective assessments and the failure to address problems caused by measurement error.

In this paper, we pursue a fifth and less widely-used Standard of Living (SoL) approach which lies somewhere between these last two approaches. The method is closely related to work on material deprivation which seeks to expand the concept of poverty beyond conventional income- or consumption-based constructs (see Berthoud *et al.*, 1993; Zaidi and Burchardt, 2005; Cullinan *et al.*, 2011). We assume that disabled people, in diverting resources to goods and services which are required because of disability, experience a lower SoL than their non-disabled counterparts. The absolute costs of disability can be identified as the additional income required by a disabled person to reach the same SoL as a non-disabled person, holding constant other characteristics, and the relative cost is the ratio of this amount to income. As Zaidi and Burchardt (2005) point out, estimates depend on the choice of a suitable standard of living indicator and the form of its relationship to income and disability status.

Our aim is to develop and improve the method further in two important respects. First, we allow for a more flexible relationship between income and SoL, so that the structure of the estimated disability cost and equivalence scale is not dictated by an unduly restrictive functional form assumption. Second, we address the problem of measurement error in disability and SoL. Both SoL and disability status are typically measured using either a binary classification or a count index based on a range of different questionnaire items.[39] Although sensitivity analyses

---

[39] The Katz activities of daily living (Katz *et al.*, 1963) and Barthel indices (Mahoney and Barthel, 1965) are two widely used tools for assessing ability to perform activities of daily living. These indices assign scores to self-reported degrees of difficulty in performing a number of activities, such as feeding, dressing, moving, bathing etc. Scores for each item are then aggregated. These indices have been criticized for the way reported difficulties are aggregated and for not taking account of potential measurement errors in self-reported difficulties (Feinstein *et al.*, 1986; Hartigan, 2007).

are often used to assess robustness, this is not effective if all the alternatives entail similar measurement error biases. To address this we use a latent factor model for disability and SoL, which explicitly allows for the existence of measurement errors in the observable indicators.

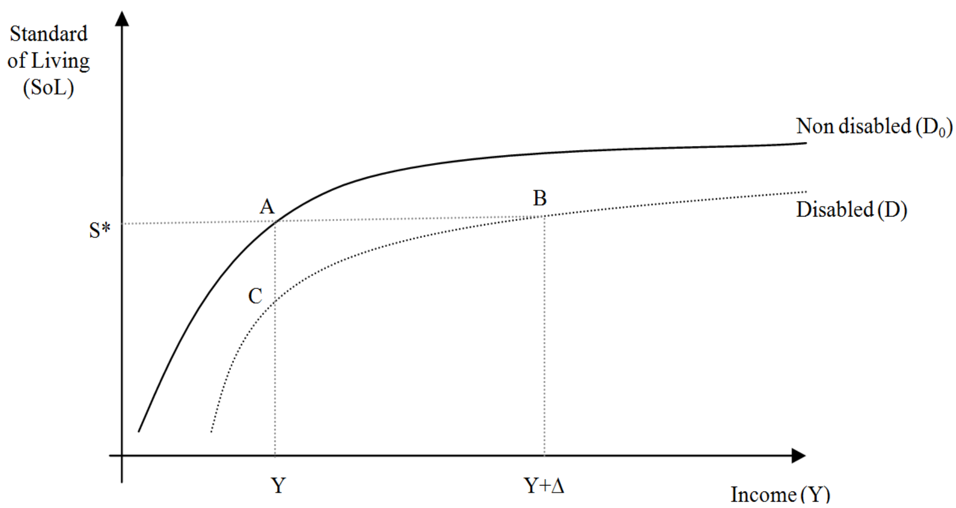
Using a two-latent factor structural equation model we estimate the extra cost of disability for a representative sample of people over state pension age living in private households in Great Britain, who were interviewed in the 2007/8 Family Resources Survey (FRS). Ten indicators of ability to afford particular items or activities are used to construct a latent continuous index of SoL. The latent SoL is modelled as a function of income, (latent) disability, and other characteristics, which reflect the many factors which determine an individual's achieved standard of living. In line with previous work (Hancock *et al.*, 2013), disability is assumed to be a latent concept which can be measured imperfectly by a vector of survey indicators reflecting difficulties in domains of life and is influenced by observed socio-economic and demographic characteristics of the individual.

This paper is organized as follow. Section 2 briefly describes the standard of living approach and its usage. Section 3 presents the latent-factor structural equation framework we employ. Section 4 describes the data used. Section 5 presents estimates of the structural equation model and derives the associated estimated extra costs of disability. Section 6 reports some sensitivity analysis on the initial results. The final section draws conclusions.

2 The Standard of Living method

Berthoud (1991) reviews various early attempts at conceptualizing and quantifying how standard of living (SoL), income and disability are related. Berthoud *et al.* (1993) and Zaidi and Burchardt (2005) formalized this approach, which has been used also by Saunders (2007) and Cullinan *et al.* (2011) for estimating the cost of disability in Australia and Ireland respectively. The SoL approach is illustrated in Figure 1, where we compare a positive level of disability D with the baseline of non disability, D_0.

FIGURE 1: Standard of Living, Income and Disability



The two curves plot the relation between income and SoL conditional on disability, and are assumed to increase monotonically with income. For any given value of income, the SoL of the disabled person lies below that of the non-disabled person and the vertical distance AC measures the difference in their standards of living at the level of income Y. This measure is similar to Sen's concept of "conversion handicap" (Doessel and Williams, 2011). The horizontal distance AB provides a measure of the extra income ($\triangle$) required to bring the SoL of the disabled person up to the same level as the non-disabled person.

To formalise this idea, consider the following additively separable SoL function:

$$S = f(Y) - g(D) + h(X, \varepsilon) \tag{1}$$

where S is the SoL, Y is a measure of financial resources, D is the degree of disability status and X and ε represent other observable and unobservable individual characteristics. Some individuals may be in receipt of disability benefit (B), others may not. To allow for this, we decompose income as:

$$Y = Y_0 + B \tag{2}$$

where Y_0 excludes disability benefits. Now define a reference level of disability D_0 and assume that the reference non-disabled person receives no disability benefit. We now pose the following question: what is the smallest amount of additional income, over and above Y_0, that would be needed for a person with disability level D to achieve the same SoL as he or she would have with income Y_0

and disability reduced to the reference level D_0? Given the additivity of (1), this additional income need, Δ, is independent of X and ε, and solves the following optimisation problem:

$$min\,\Delta \quad subject\ to\colon\ f(Y_0 + \Delta) - g(D) \geq\ f(Y_0) - g(D_0) \tag{3}$$

In general, the total disability-induced living cost Δ and the associated proportional equivalence scale $\sigma = (Y_0 + \Delta)/Y_0$ depend on the levels of both income Y_0 and disability.

For the cost Δ to depend only on severity of disability D (as implied by the design of some benefit systems), the income-SoL profile must have the linear form $f(Y_0) = \gamma_1 Y_0$, in which case the cost of disability and associated equivalence scale are:

$$\Delta = \frac{g(D) - g(D_0)}{\gamma_1}; \qquad \sigma\ = 1\ + \frac{g(D) - g(D_0)}{f(Y_0)}. \tag{4}$$

For the equivalence scale σ to depend only on disability would require $f(Y_0 + \Delta) = f(\sigma Y_0)$ to be expressible as $f(Y_0) + a(\sigma)$, for all positive σ and some function $a(.)$. The only function satisfying this property is $f(Y_0) = \gamma_1 \ln(Y_0)$, which implies the following cost of disability and equivalence scale:[40]

$$\Delta = Y_0 \left[e^{\frac{g(D)-g(D_0)}{\gamma_1}}\ -\ 1\right]\ ; \qquad \sigma = e^{\frac{g(D)-g(D_0)}{\gamma_1}}. \tag{5}$$

[40] Strictly speaking, f can be any affine transform of $\ln(Y_0)$; but an additive translation has no effect.

This is the form usually adopted for equivalence scales designed to adjust for demographic differences between households in conventional income inequality analysis. Both the linear and log-linear specifications have the advantage of simplicity and incorporate the property of base independence (or invariance of the equivalence scale to income level) in additive or multiplicative form (Lewbel, 1997).

In addition to these standard forms, we also use a more flexible log-quadratic function of the kind that has been found useful in Engel curve studies (Banks *et al.*, 1997) and embodies the constant-σ model as a special case. If $f(Y_0)$ is specified as:

$$f(Y_0) = \gamma_1 \ln(Y_0) + \gamma_2 [\ln(Y_0)]^2 \tag{6}$$

then the solution to (3) gives the cost of disability and equivalence scale as:

$$\Delta = exp\left[\frac{-\gamma_1 - sgn(\gamma_2)\sqrt{{\gamma_1}^2 - 4\gamma_2 C}}{2\gamma_2}\right] - Y_0 \tag{7}$$

$$\sigma = {Y_0}^{-1}\, exp\left[\frac{-\gamma_1 - sgn(\gamma_2)\sqrt{{\gamma_1}^2 - 4\gamma_2 C}}{2\gamma_2}\right] \tag{8}$$

where $C = -[\gamma_1 \ln(Y_0) + \gamma_2[\ln(Y_0)]^2 + g(D) - g(D_0)]$. Note that this solution requires the condition $C \leq {\gamma_1}^2/4\gamma_2$ to be satisfied.

This emphasises the importance of the specification used to relate SoL to income and the need to allow for the possibility of departures from the simple assumptions of linear or log-linear forms.

3 A statistical model

We use the following two-latent factor simultaneous equation model:

$$S_{iq} = \mathbf{1}(\lambda_q \varphi_i + \zeta_{iq}) \tag{9}$$

$$D_{ik} = \mathbf{1}(\mu_k \eta_i + \xi_{ik}) \tag{10}$$

$$\varphi_i = f(Y_i; \boldsymbol{\gamma}) + \alpha_1 \eta_i + \boldsymbol{\alpha}_2 \boldsymbol{x}_i + \varepsilon_{1i} \tag{11}$$

$$\eta_i = \boldsymbol{\beta} \boldsymbol{z}_i + \varepsilon_{2i} \tag{12}$$

where i denotes sampled individuals ($i = 1 \ldots N$), $f(.)$ represents the linear, log-linear or log-quadratic function and $\boldsymbol{\gamma}$ contains the corresponding coefficients. The latent measure of SoL is φ_i which underlies the observed SoL indicators $S_{i1} \ldots S_{iQ}$, and the latent disability index $\boldsymbol{\eta}_i$ generates observed disability indicators $D_{i1} \ldots D_{iK}$. The parameters λ_q and $\boldsymbol{\mu}_k$ are factor loadings associated with the S_{iq} and D_{ik} indicators respectively. ζ_{iq} and ξ_{ik} are the measurement errors associated with the SoL and disability indicators. The indicator function $\mathbf{1}(.)$ maps the latent indices on the right-hand side of the measurement equations (9) and (10) into the observed binary indicators of SoL and disability.

Observable covariates representing personal characteristics and household circumstances appear in vectors $\boldsymbol{x}_i$ and $\boldsymbol{z}_i$. They contain socio-economic and demographic influences on living standards and disability respectively. In this model socio-economic factors have both a direct and an indirect effect on SoL. Income, for example, has the direct effect of increasing resources available for consumption; this is captured by the function $f(Y_i;\boldsymbol{\gamma})$. Income also has an indirect influence on disability, through the term $\boldsymbol{\beta z}_i$, which then increases disability-related costs through the term $\alpha_1\eta_i$. The use of a latent disability model allows us to separate these direct and indirect effects. Note that the income concepts relevant to the direct and indirect paths are different. The direct effect involves current resources available for consumption, which includes receipt of disability benefit. In contrast, modeling of the indirect effect requires a long-term concept of economic resources reflecting the cumulative effect of past living standards on the current health state. Since disability precedes the receipt of disability benefit, it follows that the latter should be excluded from the income variable used to capture the indirect causal path.

We use the standard normalisations $corr(\varepsilon_1,\varepsilon_2)=0$ and $var(\varepsilon_1)=var(\varepsilon_2)=1$ for the structural errors and assume the measurement errors ζ_{iq} and ξ_{ik} to be independent. Because units of measurement for φ and η are arbitrary, we show coefficient estimates in standardized form. The variance of the latent SoL index in (11) is $(1-R^2_\varphi)$, where $R_{\boldsymbol{\varphi}}{}^2$ is the squared multiple correlation of φ, so the standardised form of $\boldsymbol{\varphi}$ implies multiplying each coefficient by a factor $(1-R^2_\varphi)^{-1/2}$, so that each coefficient is interpretable as the change in $\boldsymbol{\phi}$ in standard deviation units, produced by a 1-unit increase in the value of the covariate.

Disability η is also a latent construct, with variance $var(\boldsymbol{\beta z}_i) + 1 = (1 - R^2_\eta)$, where R^2_η is the squared multiple correlation of the disability equation. Therefore the standardized coefficient of φ on η is $\alpha_1^{STD} = \alpha_1\sqrt{[(1-R^2_\eta)/(1-R^2_\varphi)]}$, which can be interpreted as the change in φ (in standard deviation units) generated by a 1-standard deviation increase in η.

4 Data

The data are from the 2007-8 Family Resources Survey (FRS): a large UK household survey collecting detailed income and assets information from respondents and asking questions covering difficulties due to ill-health or disability. The survey also includes a series of questions aimed at measuring material deprivation (Department for Work and Pensions, 2009). For this paper we restrict the analysis to households in Great Britain where all members are aged over state pension age (65 for men; 60 for women) and the household contains only a single person or a couple. The age restriction is imposed in order to limit endogeneity bias which may arise for younger adults for whom disability may cause a reduced income by limiting labour market participation.[41] In estimating equations (11) and (12) we measure income at the household level assuming that all members of the households benefit to the same extent from total household income. This

[41] For a discussion on this point we refer, amongst others, to Goldman (2001) and Adams *et al.* (2003).

is less likely to be true for households containing members other than a single or couple pensioner. After dropping a few cases where relevant information is missing the resulting sample contains 8,183 individuals (5,812 households). About 58% of the sample are partnered and the remainder live alone. We retain proxy cases (4.8%) where the required data was provided by a proxy respondent (often a carer). Dropping proxy cases would bias the sample towards the less severely disabled.

Deprivation indicators are derived from a set of questions about items or activities, seen as potential 'necessities'; households who did not have the items or do the activities were asked whether this was because they did not want them or because they could not afford them. They were also given the option of saying that an item or activity did not apply to them.[42] From these household-level indicators, we created individual-level indicators in which each household member is assigned the values of the deprivation indicators of their household. Each indicator is set to 1 if the respondent answered "*We/I would like to have this but cannot afford this at the moment*" and 0 otherwise. Thus we allow for differences in preferences to explain non consumption rather than assuming that non consumption always implies deprivation. However, it has been suggested (McKay, 2004, 2008; Berthoud *et al.*, 2009) that certain segments of the population with

[42] Taking the ability to afford to replace worn out furniture as an example, respondents who rent furnished properties may not be responsible for replacing furniture and therefore select 'does not apply'. In fact only 2.5% of the sample replied 'does not apply' to at least one of the deprivation indicators.

lowered expectations, such as disabled older people, may be less likely than others to admit to being unable to afford particular activities or goods. We have carried out two sensitivity analyses by: (*i*) using a restricted subset of the indicators; and (*ii*) using a less stringent interpretation of the responses. Results are given in section 6 below.

In estimating equations (9)-(12), we invert these deprivation indicators to construct the SoL indicators, S_{iq}, taking the value 0 if the respondent cannot afford the activity/good and 1 otherwise. Sample statistics corresponding to the two alternative definitions of deprivation are shown in Appendix Table A1. Overall, 35% of the sample report an inability to afford at least one item, a proportion which rises to 80% under the less stringent interpretation.

FRS respondents are asked whether they have a health problem or disability and, if they answer 'yes' they are asked if they have significant difficulties in each of nine areas of life. The prevalence rates for these disability indicators are reported in Appendix Table A1. Overall, 53% of the sample reported having no disability and 20% reported three or more difficulties. The most common difficulties are those concerning physical impairment (difficulties in mobility; with lifting, carrying or moving objects).

The explanatory covariates used in the SoL and disability equations are summarised in Appendix Table A2. The income indicator Y used in the SoL equation represents the resources of the household currently available for meeting the consumption needs of the household members. We use a household-level income measure, net of direct taxes and housing costs, similar to the “After Housing Cost” measure used in the official *Households Below Average Income* analysis (DWP, 2009) and also by Zaidi and Burchardt (2005). This measure represents the disposable income available for spending on the items and activities used as indicators of SoL. We argue that the treatment of housing as a fixed cost is reasonable in our target population, since adjustment of housing as a response to disability often takes the form of transition into the care home sector or moving into a multi-generation household. Nevertheless, we report a sensitivity test in section 6 below.

Our income measure includes income from investments (interest, rent, dividends, private pensions, annuities). It includes disability benefits since, as argued earlier, they are available, like any other income component, to be used to maintain SoL (see also Zaidi and Burchardt, 2005; Stapleton, 2008 and Cullinan *et al.*, 2011). Disability benefits comprise the non-means-tested Attendance Allowance and Disability Living Allowance, an estimate of income attributable to the Severe Disability Premium component of means-tested pensioner benefits and other minor disability-related benefits that are received by a small number of older people in our sample.

The income measure used as a covariate in the disability equation also includes income from investments, since interest, rent, dividends, private pensions and annuities are returns on assets accumulated over the lifecycle and are, consequently, good indicators of past access to resources with a cumulative positive influence on health. For the same reason, we also include a measure of financial wealth[43] in the disability equation and a dummy variable to indicate home ownership. Note that the income measure used as a covariate in the disability equation excludes current receipt of disability benefits, since those are a consequence, rather than a determinant, of current disability. Rather than use an arbitrary equivalence scale to adjust income for household composition, we include a dummy variable to indicate whether the household contains a single person or a couple in the disability and SoL equations. In line with previous work (Zaidi and Burchardt, 2005; Stewart, 2009), we also use a set of personal characteristics including age, gender, level of education, home ownership and marital status, together with regional dummies to reflect geographical differences in cost of living and in health.

[43] Deposit and saving account balances, stocks, bonds, certificate deposits and other savings held by the household. The information recording the amount of liquid wealth in FRS was severely affected by non-response, which we deal with by imputation based on grossing up investment income. Financial wealth is not used as a covariate in the SoL equation.

5 Parameter Estimates and Analysis

5.1 Estimates of the Structural Equation Model

Estimation results for the model comprising equations (9)-(12) are presented in Appendix Tables A3-A5.[44] The log-quadratic form of the SoL equation fits the data best. The estimated measurement equations (9) and (10) using this form of the SoL equation, are summarised in Appendix Table A3. They show respectively the factor loadings λ_q which capture the effect of the latent standard of living index φ on the indicators S_q, and the factor loadings μ_k associated with the disability score η. We also report the squared correlation of each indicator with the underlying latent construct. The factor loadings are all positive and highly significant. Being unable to afford to replace/renew durable goods or to keep the home in a decent state of decoration are the most sensitive indicators of the latent SoL construct φ; the inability to afford house insurance, hobbies or leisure activities are the least sensitive. The highest correlation with the latent disability construct is found for indicators of difficulties with mobility, lifting and dexterity, while lower correlations are found for indicators of cognitive disability.

Results reported in Appendix Table A4 show that the conditional mean of η increases almost linearly with age, although we allowed for non-linearity using a spline function of age, with a single node at the median age 73 observed in the

[44] Estimates were computed using the robust maximum likelihood estimator of *Mplus 6.11* (Muthén and Muthén, 2010).

sample. The structural estimates provide no evidence of a significant relation with gender. Indicators measuring economic well-being are jointly significant at the 1% level: more educated individuals experienced a low level of disability as well as those with high current pre-disability benefit income. A negative relation between wealth and disability emerges, both in terms of housing wealth (captured by owner-occupation) and financial wealth.

Income and receipt of disability benefits by decile of latent disability are displayed in Table 1. Average weekly post-disability benefit household income (Y) is reported per-capita and without adjustment for household composition. The association between disability and socio-economic status is widely recognized (see for instance Cutler *et al.*, 2011 and Goldman, 2001 for a review) although the extent to which this association reflects causality is still in debate (Conti *et al.*, 2010). Similarly we find that there is a strong association between disability and per-capita income which declines monotonically until the fifth decile of η and is almost flat afterwards. Thus poor health and low income are strongly associated even if the measure of income used, as here, includes the disability benefit that individuals receive. The last three columns of Table 1 show the percentage of individuals in the sample in receipt of any disability benefit by decile of latent disability, the proportions of those recipients who are in each disability decile and the average amount of disability benefits received by individuals in each disability decile. The proportion of individuals in the sample who receive these benefits ranges from under 2% in the lowest disability decile to 50% in the top decile. Overall, amongst those in the upper half of the disability distribution the percentage is 27%. Although current disability benefits appear well targeted on disabled people, a significant proportion of those who face severe disability do not receive disability benefits. Non take-up of disability benefits among disabled people has been noted elsewhere (Pudney, 2010, Currie and Madrian, 1999) and the receipt of disability benefit may often be delayed by several years after disability onset (Zantomio, 2013).

Table 1: Mean income and receipt of disability benefits by deciles of latent disability

Decile of $\hat{\eta}$	Mean Y[a] £s pw		% of individuals receiving disability benefits	% of individual disability benefit recipients in each disability decile	Average amount of disability benefit[b] received (£s pw)
	Per capita	Unadjusted for household composition			
1	263.90	442.90	1.8	1.2	1.20
2	206.00	353.10	3.2	2.0	1.90
3	187.40	309.10	3.6	2.3	2.30
4	162.80	257.70	5.0	3.2	2.70
5	141.30	203.80	6.9	4.4	4.00
6	148.70	221.50	10.3	6.6	6.30
7	172.20	264.10	15.5	10.0	10.00
8	175.50	263.80	24.1	15.4	15.70
9	174.10	255.50	35.6	22.8	23.50
10	181.70	264.10	50.1	32.1	37.80
Mean for deciles 6 to 10	*170.40*	*253.80*	*27.1*	*86.9*	*18.60*

Notes: Statistics computed over a sample of 8,183 FRS 2007-8 respondents. All monetary values are rounded to the nearest 10p and expressed in 2007 prices.
a. Household income including disability benefit.
b. Measured at the individual level.

Estimates for the regression coefficients of the SoL equation are reported in Appendix Table A5, using three different functional forms of $f(Y)$: the linear-in-income (model 1); the linear-in-log income (model 2); and the quadratic-in-log income model (model 3). Age, level of education, home ownership, marital status and region of residence are found to be highly significant at the 1% level and their signs, for the most part, are as expected. A gender dummy is not significant. Here, we focus on the structural parameters of interest in deriving the equivalence scale (Table 2). The structural estimates of the α_1 and $\boldsymbol{\gamma}$ provide strong evidence that latent disability and current income affect the SoL. Increased disability is associated with lower values of the SoL index, while income is positively associated with the SoL, no matter which functional form is used. Holding other variables constant, a 1-standard deviation increase in disability η produces a reduction of 0.233 standard deviations in φ using model 1, 0.254 using model 2 and 0.236 using model 3. The estimated income coefficients imply that a £10 increase in weekly income increases SoL by 0.03 standard deviations in model 1 and in model 2; a 10% increase in net income produces an increase of about 0.0631 standard deviations in the SoL. In model 3, the coefficient associated with the added square of log household income is significant at the 1%, implying a significant non-linear relationship of income and the SoL index φ. Thus, controlling for disability level, disability costs appear to vary with income in both absolute terms and as a proportion of income.

At the bottom of Table 2 we report the number of free estimated regression parameters, the maximized log-likelihood and its correction for non normality factor, the Akaike information criterion AIC and the Bayesian Information criterion BIC for the model comprising equations (9)-(12). According to these measures the quadratic-in-log form (model 3) fits the data best but, as the plots in Appendix Figure A1 show, its implications are remarkably close to those of the linear specification.

We might also want to include covariates in the SoL equation which capture the value of any informal (i.e. unpaid for) and subsidised formal care received by the person, as such care may affect the living standard a disabled person can achieve from a given level of income. Informal care received by another member of the household can be ignored as it represents a within-household transfer rather than an addition to household resources. The FRS contains limited information on receipt of informal care from non-household members and formal care although whether and how much that care was subsidised by the state is not directly recorded. We experimented with adding covariates for hours of informal care received from non-household members and hours of care from a Local Authority or nurse, in the SoL equation (income entered in log-quadratic form). None of the estimated coefficients was statistically significant at the 5% level and the estimated coefficients for latent disability and income were only very marginally changed by the inclusion of these additional covariates. In subsequent analysis we therefore use the models without covariates measuring receipt of care.

Table 2: The standard of living equation: parameter estimates (in standardised form) for latent disability and income

	Model (1) linear in Y		Model (2) linear in ln(Y)		Model (3) quadratic in ln(Y)	
Parameter(s):	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
α_1^{STD}	-0.233***	0.016	-0.254***	0.016	-0.236***	0.016
γ_1^{STD}	0.003***	0.001	0.631***	0.026	-2.610***	0.201
γ_2^{STD}					0.307***	0.019
Free parameters	74		74		75	
Log-likelihood	-38718.413		-38759.401		-38694.623	
Correction for non-normality factor	1.004		0.992		0.994	
AIC	77584.826		77666.803		77539.247	
BIC	78103.552		78185.529		78064.983	

Notes: Significance: * = 10%; ** = 5%, *** = 1%. Models also include regional dummy variables and controls for socio-economic characteristics which are reported in Appendix Table A5. The R^2 of model (1), (2) and (3) are 0.384; 0.334; and 0.382, respectively.

5.2 Disability Costs and Equivalence Scales

Using the parameter estimates in Table 2, we can derive the relative/absolute costs of disability for any reference level of disability D_0 as the minimal compensating amount (3). First, we calculate the model-based posterior prediction $\hat{\eta}$ as the estimate of the expectation of η conditional on all observed information for the individual. Then we calculate the estimate of disability cost as (4), (5) or (7) evaluated at the point $\hat{\eta}$ and thus the means of these estimated costs by decile of $\hat{\eta}$.[45]

Since we use a continuous measure of disability, the definition of D_0 is less straightforward than when using a dichotomous indicator. We can think of D_0 as a reference level of disability above which some financial compensation is judged appropriate, but how should this reference level be chosen? Table 3 reports the prevalence of reported difficulties by decile of $\hat{\eta}$. As noted in section 4, about 53% of the sample reported having no disability. All individuals who fall in the highest four deciles of $\hat{\eta}$ reported at least one disability, most having a difficulty with mobility, lifting, carrying or moving objects. The mean number of reported disabilities increases non-linearly with position in the latent disability distribution. It is clear from Table 3 that there is a definite discontinuity at the median and, as a consequence, we adopt the median level of $\hat{\eta}$ ($D_0 = 0.972$) as our reference level. Appendix Figure A2 shows the empirical kernel distribution of the predicted disability index $\hat{\eta}$ from the log-quadratic model.

Estimated costs of disability are presented in Table 4. There are 260 cases (out of 8,183 in the estimation sample) where the condition $C \leq {\gamma_1}^2/4\gamma_2$ in equation (8) is violated. All have a combination of low income (mean £88 compared to

[45] Note that this is a conservative estimate, for the log-linear and (to a lesser extent) the log-quadratic model. Because of the convexity of the $exp(.)$ function in (5) and (7), the true average cost will be understated: to a degree that depends on the posterior variance of η.

£290 for the full sample) and low estimated latent disability (mean 0.65 compared to 1.40). In the calculations reported below, we set their disability costs to zero (dropping cases with very low income and disability leads to virtually identical estimates).

Table 3: Self-reported difficulties by decile of $\hat{\eta}$

Decile of $\hat{\eta}$	% of those who reported any difficulties	difficulties with mobility, lifting, carrying or moving objects	Number of difficulties reported
1	0.0%	0.0%	0.00
2	0.0%	0.0%	0.00
3	0.0%	0.0%	0.00
4	0.2%	0.0%	0.00
5	5.5%	0.0%	0.06
6	63.2%	2.3%	0.67
7	100.0%	91.8%	1.22
8	100.0%	98.4%	2.22
9	100.0%	100.0%	3.03
10	100.0%	100.0%	4.97
Mean	46.9%	39.2%	1.22

Notes: Statistics computed over a sample of 8,183 FRS 2007-8 respondents.

Average estimated disability costs (Δ) and the equivalence scale (σ) computed among people upper the 50% of the disability distribution are displayed in Table 4 by deciles of $\hat{\eta}$ (panel *a*) and by deciles of household income (panel *b*) for each of the three model variants. From panel *a)*, we see that on average, a person in the upper 50% of the disability distribution requires an additional £90 to reach the same standard of living as a comparable person at the median level of disability, according the linear model. Average disability costs are about £17 per week in the sixth decile of the disability distribution, rising to £164 in the top decile. For the log-linear specification the estimated disability costs are higher (about £154 per week for those in the upper 50% of disability) and they increase more sharply with disability. The log-quadratic model generates estimates which are much closer to those of the linear model, but with slightly higher values in the upper tail of the disability distribution. The estimated average cost of disability among the upper 50% of disabled people is about £99 per week; in the top decile of disability it is £180. Panel *b)* of Table 4 reports equivalence scales and disability costs among disabled people by deciles of per capita pre-disability benefit income. It demonstrates that the flexible log-quadratic model allows for a more complex relationship between income and estimated disability costs/equivalence scales than the other two models. Under the log-quadratic model the estimated costs of disability are greatest for the lowest and highest income decile. The estimated equivalence scale is largest for the lowest income decile.

Table 4: Estimated costs of disability and average equivalence scale among disabled people[a], by deciles of latent disability and by deciles of per capita income[b]

Panel *a)*

Decile of latent disability, $\hat{\eta}$	Model (1) linear in Y		Model (2) linear in ln(Y)		Model (3) quadratic in ln(Y)	
	Δ £s pw	σ	Δ £s pw	σ	Δ £s pw	σ
6	17.40	1.11	23.10	1.10	22.10	1.21
7	62.00	1.35	95.10	1.38	67.80	1.40
8	91.00	1.50	149.60	1.60	98.00	1.54
9	116.30	1.72	193.10	1.83	126.10	1.78
10	163.70	2.06	307.50	2.36	179.90	2.17
Mean among disabled people	*90.00*	*1.55*	*153.60*	*1.65*	*98.70*	*1.62*

Panel *b)*

Decile of per capita pre-disability benefit income *(% of disabled people in each decile)*	mean Y[c] £s pw		Model (1) linear in Y		Model (2) linear in ln(Y)		Model (3) quadratic in ln(Y)	
	Per-capita	Unadjusted for household composition	Δ £s pw	σ	Δ £s pw	σ	Δ £s pw	σ
1 *(60.1%)*	95.10	141.50	95.60	2.16	81.50	1.72	115.60	2.65
2 *(59.6%)*	109.40	180.20	88.10	1.63	106.10	1.64	89.70	1.65
3 *(56.7%)*	127.60	204.50	89.80	1.56	120.70	1.65	90.90	1.57
4 *(63.8%)*	136.20	186.60	87.30	1.58	111.30	1.63	87.90	1.57
5 *(55.0%)*	149.30	211.40	89.20	1.51	133.20	1.65	91.00	1.51
6 *(52.1%)*	167.90	249.50	92.70	1.47	155.00	1.68	96.20	1.47
7 *(46.2%)*	192.90	287.40	89.90	1.39	167.40	1.65	95.40	1.41
8 *(42.7%)*	226.40	344.70	88.80	1.31	205.80	1.63	100.80	1.34
9 *(33.2%)*	270.00	403.50	90.10	1.27	246.90	1.64	108.60	1.31
10 *(30.4%)*	412.10	594.90	88.10	1.17	377.80	1.63	130.20	1.23
Mean among disabled people	*170.40*	*253.80*	*90.00*	*1.55*	*153.60*	*1.65*	*98.70*	*1.62*

Notes:

a. Disabled people defined as those in the upper 50% (deciles 6-10) of the distribution of disability, $\hat{\eta}$.

b. Income deciles computed over the pre-disability benefits income distribution of the whole population which includes non-disabled people.

c. Y is household income including disability benefit.

Estimates of Δ are unadjusted for household composition. All monetary values are rounded to the nearest 10p and expressed in 2007 prices. The reference disability level for computing Δ and σ is the median.

It is clear that the equivalence scale, σ increases with disability.[46] If we define a disabled person as someone with a disability in the top half of the disability

[46] By construction,   obtained using model 1 and 2 is lower than 1 for those individuals who fall below the median level of disability $[g(D) < g(D_0)]$ and increases afterwards. However, nothing prevents the equivalence scale derived from model 3 for some people with disability level below D_0 from being greater than 1. That is because the equivalence scale derived from specification 3, while increasing in disability, is decreasing in income. In practice this occurs for only 1.07% of the sample.

distribution, an older disabled person requires, on average, an increase of about 55% of net weekly pre-disability household income (Y_0) to reach the same standard of living as a comparable non-disabled person, according to the linear model. Average disability costs are about 11% of Y_0 in the sixth decile of the disability distribution, rising to 106% in the top decile. For the log-linear specification, estimated disability costs are about 65% higher on average in the disabled population and increase more sharply with disability. The log-quadratic model generates estimates which are much closer to those of the log-linear model, but with slightly lower values in the upper tail of the disability distribution. The average extra cost of disability is about 62% of the net weekly pre-disability household income.

6 Sensitivity analysis

In this section we assess the sensitivity of our results to: (*i*) the assumption that the costs of disability and the equivalence scale are independent of household composition; (*ii*) the income definition; and (*iii*) the construction of the SoL measure.

Demographic invariance. The three models of the previous section imply invariance of the equivalence scale to household size and structure. This has the advantage that a benefit system with the same property does not create incentives for potential claimants to change their household type to increase their level of entitlement (Pendakur, 1999). We test whether estimates of the best-fitting quadratic model are sensitive to the assumption of demographic invariance by using a two-group analysis where we allow the parameters of the SoL equations (9) and (11) to differ for respondents from single-person and two-person households. In contrasting this with the unrestricted model, the Akaike information criterion suggests that the unrestricted model provides a slightly better balance of model fit and parsimony. Panel (1) in Table 5 shows the equivalence scale and the extra cost of disability computed for single people and couples, by disability index η. It should be noticed however, that about 58% of single people, compared with 44% of couples, belong to the top four deciles of $\hat{\eta}$. Thus single people (mainly widows) on average experience higher disability levels than people in couples (see also Zaidi and Burchardt, 2005). On the other hand, household income (not adjusted for household composition) of people in couples is generally higher than for single people. So that the reduction in the living standard caused by a given disability level is higher (lower) in relative (absolute) terms for single people than couples.

Housing wealth and housing costs. A further sensitivity analysis makes some allowance for housing wealth. We re-estimate equations (9)-(12) adding to the income variables in equations (11) and (12) an annual return from the (estimated) house wealth of 2% and 4%, respectively.[47] This increases the household income measure only for the 76% of people who are owner occupiers. Estimates of equivalence scales and the extra costs of disability using a 2% and 4% return on housing wealth are remarkably close to the base case and are reported in panel (2) of Table 5. We also test the extent to which our estimates are sensitive to the treatment of housing costs in the income measure. On average, housing costs (which are the sum of gross rents, council tax payments, costs of insurance on structure of property and mortgage interest payments net of housing benefit and council tax benefits) are of about £8 lower for disabled people compared with the non-disabled counterpart. Using a "Before Housing Costs" income measure (see discussion in section 4) yields an estimate of the average extra cost of disability among disabled people of £93 (about £6 lower than when income is measured after housing costs).

SoL indicators. We used two sensitivity tests focused on disability measurement. First, dropping the indicators for "hobby or leisure activity", "holidays away from home" and "friends and family round" produced very little change in

[47] Estimates of housing wealth are derived by estimating an interval regression using recorded Council Tax band information and a set of controlling characteristics available in the FRS. Council Tax is a local property tax for which all domestic properties have been valued and the value placed in a band. This regression gives us a vector of estimated coefficients with we use to derive homeowners' expected housing wealth conditional on being in the respondent council tax band, evaluated at the time when their properties were last valued (1991 for England and Scotland and 2005 in Wales). Finally, observed regional changes in house prices between then and 2007 are applied to yield estimated housing wealth in 2007 prices. Return on housings wealth is then computed at a weekly basis (dividing the assumed annual return by 52).

the estimates.[48] Second, we used a less stringent interpretation, setting each indicator to 0 even in cases where respondents replied *"We/I do not want/need this"*. This produced a slightly lower coefficient (-0.272) for disability, a higher γ_1^{STD} (-1.793) and lower γ_2^{STD} (0.217), yielding an estimate of the extra cost of disability among disabled people of about 89% of their household income. Results are shown in panel (4) of Table 5.

Table 5: Sensitivity Analysis: mean costs of disability and equivalence scale by deciles of η

Decile of $\hat{\eta}$	(1) Couples (N=4,752)		(1) Singles (N=3,438)		(2) Returns from Housing wealth 2%		(2) Returns from Housing wealth 4%		(3) SoL indicator = 0 if does not want/ have / or cannot afford to, 1 otherwise	
	Δ £s pw	σ	Δ £s pw	σ	Δ £s pw	σ	Δ £s pw	σ	Δ £s pw	σ
6	23.60	1.11	14.70	1.15	21.40	1.20	21.70	1.20	29.90	1.23
7	74.00	1.31	55.40	1.42	66.30	1.39	67.30	1.40	100.70	1.56
8	108.20	1.43	80.60	1.57	96.10	1.52	97.40	1.53	148.90	1.78
9	139.00	1.55	102.40	1.82	123.50	1.76	125.20	1.77	192.20	2.13
10	197.10	1.82	147.40	2.25	176.30	2.14	178.90	2.15	281.30	2.74
Mean for deciles 6 to 10	*107.60*	*1.44*	*80.60*	*1.65*	*96.70*	*1.60*	*98.10*	*1.61*	*150.50*	*1.89*

Notes: Estimates of Δ are unadjusted for household composition. All monetary values are rounded to the nearest 10p and expressed in 2007 prices.

7 Discussion and Conclusions

In this paper, we have applied the standard of living approach to estimate the cost of disability among older people in Great Britain and extended previous

[48] We estimated α_1^{STD}, γ_1^{STD} and γ_2^{STD} as -0.236, -2.660 and 0.311 respectively, compared with -0.236, -2.610 and 0.307 for the baseline model.

research by developing a two-latent factor structural model to estimate equivalence scales for disability. Disability is treated as a latent construct which is measured imperfectly by a vector of survey indicators and is influenced by observed socio-economic characteristics. Ten indicators of deprivation are used as observable counterparts of the latent continuous index of SoL, which varies in relation to household income and disability. Our approach allows us to construct a base-dependent equivalence scale (i.e. one which varies by income level) which takes account of the severity of disability. The restrictions on preferences imposed by the assumption of a base-independent equivalence scale for disability are not supported by our data. This implies that the extra income that disabled people on higher incomes need to be as well off as their non-disabled counterparts is lower than the equivalent sum needed by disabled people on lower incomes. Our application is the first, in our knowledge, to derive an equivalence scale for disability using a log-quadratic function on income of the kind that has been used in Engel curve studies.

The results show that the extra costs of disability are substantial, and rise with severity. Using the 2007/8 wave of the FRS we estimate that an older disabled person, defined as someone above the median level of disability for all older people, requires a net household income around 62% higher than that of a comparable person with a median level of disability to reach the same standard of living. This corresponds to around £99 per week on average as an allowance for the additional costs that households with a disabled member face. These additional costs where disability is in the highest decile of disability average £180 under our preferred model. The latter is comparable with disability costs for highly disabled pensioners estimated by Zaidi and Burchardt (2005) which ranged from £122 to £190 (converted to 2007 prices from £104 to £162 in 2002 prices).

Only about 27% of those whom we estimate to face disability-related costs, are in receipt of disability-related cash benefits. In line with previous findings (Berthoud *et al.*, 1993; Thompson *et al.*, 1990) we find evidence that, although disability benefits are received mainly by people who do indeed face disability costs, they do not meet the full costs of disability for recipients, and a high proportion of people with severe disability do not receive disability benefits at all.

We have also investigated the sensitivity of our estimates to various aspects of the econometric specification, the measurement of SoL and the treatment of housing wealth and costs. Estimates obtained using the preferred quadratic model are remarkably close to those obtained when a simple linear-in-income form is used. The estimates are sensitive to whether the disability costs and equivalence scales are constrained to be the same for single people and couples: the reduction in living standards for a given disability level appears to be higher (but not parallel) for single people than for couples. This is in contrast to Zaidi and Burchardt (2005) who found that disability costs were higher for single people than for couples. As a consequence there is more divergence between the our and their estimates when single people and couples are distinguished. Zaidi and Burchardt found that highly disabled single pensioners faced extra costs of around £189 (2007 prices) compared with our estimate for single pensioners in the highest decile of disability of £147. The equivalent comparison for couples is £122 against our higher figure of £197. Thus while there is evidence that disability benefits systems should discriminate between single people and couples, more research is needed before firm recommendations for policy can be made. Our estimates are only marginally sensitive to the inclusion of the return on housing wealth in income.

The estimated equivalence scale is very sensitive to the way answers to survey questions on deprivation are interpreted. If we were to interpret all cases of non-possession as equivalent to deprivation, we would estimate that an older disabled person requires a net household income around 89% higher than a comparable non-disabled person to reach the same standard of living, compared with 62% when the index is based only on explicit inability to afford.

Our clear – and robust – conclusion is that disability costs faced by older people in Britain are large and increase strongly with severity of disability. Comparisons of the incomes of disabled and non-disabled older people must make adequate allowance for these costs if meaningful inferences about their relative living standards are to be drawn.

References

Adams, P., M.D. Hurd, D. McFadden, A. Merrill and T. Ribeiro, “Healthy, wealthy, and wise? Tests for direct causal paths between health and socioeconomic status,” *Journal of Econometrics*, 112(1), 3-56, 2003.

Banks, J., R. Blundell and A. Lewbel, “Quadratic Engel curves and consumer demand,” *The Review of Economics and Statistics*, 79(4), 527-539, 1997.

Berthoud, R., J. Lakey, and S. McKay, *The Economic Problems of Disabled People*, Policy Studies Institute, London, 1993.

Berthoud, R., “Meeting the costs of disability,” in G. Dalley, ed., *Disability and Social Policy*, Policy Studies Institute, London, 1991.

Berthoud, R., M. Blekesaune and R. Hancock, “Ageing, income and living standards: evidence from the British Household Panel Survey,” *Ageing and Society*, 29(7), 1105-22, 2009.

Burchardt, T., “Capabilities and disability: the capabilities framework and the social model of disability,” *Disability and Society*, 19(7), 735-751, 2004.

Conti, G., J. Heckman and S. Urzua, “The education-health gradient,” *American Economic Review*, 100(2), 234-38, 2010.

Coulter, F.A.E., F.A. Cowell and S.P. Jenkins, ”Differences in needs and assessment of income distributions,” *Bulletin of Economic Research*, 44(2), 77-124, 1992.

Cullinan, J., B. Gannon and S. Lyons, “Estimating the extra cost of living for people with disabilities,” *Health Economics*, 20 (5), 582-599, 2011.

Currie, J. and B. C. Madrian, “Health, health insurance and the labor market,” in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics, vol. 3C,* 3309-3416, Elsevier Science, New York, 1999.

Cutler, D.M., A. Lleras-Muney and T. Vogl, “Socioeconomic status and health: dimensions and mechanisms,” in S. Glied and P.C. Smith, eds., *Oxford Handbook of Health Economics*, Oxford University Press, Oxford, 2011.

Da Roit, B. and B. Le Bihan, "Similar and Yet So Different: Cash-for-Care in Six European Countries' Long-Term Care Policies," *The Milbank Quarterly,* 88(3), 286-309, 2010.

Deaton, A. and C. Paxson, "Economies of scale, household size, and the demand for food," *Journal of Political Economy*, 106(5), 897-930, 1998.

Department for Work and Pensions, *Households Below Average Income 1994/95-2007/2008*, The Stationery Office, London, 2009.

——— , *Family Resources Survey, 2007–8*, The Stationery Office, London, 2009.

Doessel, D.P. and R.F.G. Williams, "Disabled people's living standards: filling a policy vacuum," *International Journal of Social Economics,* 38(4), 341-357, 2011.

Feinstein, A.R., B.R. Josephy and C.K. Wells, "Scientific and clinical problems in indices of functional disability," *Annals of Internal Medicine*, 105(3), 413-420, 1986.

Gleckman, H., *Long-term care financing reform: lessons from the US and abroad*, Commonwealth Fund Publication, 1368, 2010.

Goldman, N., "Inequalities in health: disentangling the underlying mechanisms," in A. Weinstein and M. Hermalin, eds., *Strengthening the Dialogue between Epidemiology and Demography*, Annals of the New York Academy of Sciences, New York, 2001.

Gordon, D., L. Adelman, K. Ashworth, J. Bradshaw, R. Levitas, S. Middleton, C. Pantazis, D. Patsios, S. Payne, P. Townsend and J. Williams, *Poverty and Social Exclusion in Britain,* Joseph Rowntree Foundation, York, 2000.

Hancock, R., and S. Pudney, "Assessing the distributional impact of reforms to disability benefits for older people in the UK: implications of alternative measures of income and disability costs," *Ageing and Society*, available on CJO2012. doi:10.1017/S0144686X1200075X, 2013.

Hancock, R., M. Morciano, S. Pudney and F. Zantomio, *Do household surveys give a coherent view of disability benefit targeting? A multi-survey latent variable analysis for the older population in Great Britain,* HEG Working Paper no. 13-03, University of East Anglia, Norwich, 2013.

Hartigan, I., "A comparative review of the Katz ADL and the Barthel Index in assessing the activities of daily living of older people," *International Journal of Older People Nursing*, *2*(3), 204-212, 2007.

Jones, A. and O. O'Donnell, "Equivalence scales and the costs of disability," *Journal of Public Economics*, *56*(2), 273-289, 1995.

Kasparova, D., A. Marsh and D. Wilkinson, *The take-up rate of Disability Living Allowance and Attendance Allowance: Feasibility study*, Department for Work and Pensions, Research Report No 442, London, 2007.

Katz, S., A.B. Ford, R.W. Moskowitz, B.A. Jackson and M.W. Jaffe, "Studies of illness in the aged," *The Journal Of The American Medical Association*, 185(12), 914-919, 1963.

Lelli, S., "Using functionings to estimate equivalence scales," *Review of Income and Wealth*, 51(2), 255-285, 2005.

Lewbel, A., "Consumer demand systems and household equivalence scales," in M.H. Pesaran and P. Schmidt, eds., *Handbook of Applied Econometrics, Volume II: Microeconomics*, Blackwell Publishers Ltd., Oxford, 1997.

Mahoney, F.I. and D.W. Barthel, "Functional evaluation: the Barthel Index," *Maryland State Medical Journal*, 14(2), 61-65, 1965.

Martin, J. and A. White, *The financial circumstances of disabled adults living in private households,* HMSO, London, 1988.

Mckay, S. "Poverty or preference: What do "Consensual Deprivation Indicators" really measure?" *Fiscal Studies*, 25(2), 201-223, 2004.

——— , *Measuring material deprivation among older people : Methodological study to revise the Family Resources Survey questions*, Working Paper no. 54, Department for Work and Pensions, London, 2008.

Muellbauer, J., "McClements on equivalence scales for children," *Journal of Public Economics*, 12, 221-231, 1979.

Muthén, L.K., and B.O. Muthén, *Chi-square difference testing using the S-B scaled chi-square*, Note on Mplus website, Los Angeles, www.statmodel.com, 2005.

——— , *Mplus User's Guide. Sixth Edition,* Muthén and Muthén, Los Angeles www.statmodel.com, 2010.

Nelson, J.A., "Household equivalence scales: Theory versus policy?," *Journal of Labor Economics*, 11(3), 471, 1993.

Pendakur, K., "Semiparametric estimates and tests of base-independent equivalence scales," *Journal of Econometrics*, 88(1), 1-40, 1999.

Pollak, R.A. and T.J. Wales, "Welfare comparisons and equivalence scales," *The American Economic Review*, 69(2), 216–221, 1979.

Pudney, S., *Disability benefits for older people: How does the UK Attendance Allowance system really work?*, ISER Working Paper no. 2010-02, University of Essex, Colchester, 2010.

Saunders, P., "The costs of disability and the incidence of poverty," *Australian Journal of Social Issues*, 42(4), 461-480, 2007.

Smith, N., S. Middleton, K. Ashton-Brooks, L. Cox and B. Dobson, *Disabled people's costs of living: more than you would think,* Joseph Rowntree Foundation, York, 2004.

Stapleton, D., A. Protik and C. Stone, *Review of international evidence on the cost of disability,* Research report No. 542, Department of Work and Pensions, London, 2008.

Stewart, M.B., "The estimation of pensioner equivalence scales using subjective data," *Review of Income and Wealth*, 55(4), 907-929, 2009.

Swartz, K., N. Miake and N. Farag, "Long-term care: Common issues and unknowns," *Journal of Policy Analysis and Management*, 31 (1), 139–152, 2012.

Thompson, P., M. Lavery and J. Curtice, *Short-changed by Disability,* Disability Income Group, London, 1990.

Welfare Reform Act, 2012 available at:

http://www.legislation.gov.uk/ukpga/2012/5/contents/enacted

Wittenberg, R., B. Hu, R. Hancock, M. Morciano, A. Comas-Herrera, J. Malley and D. King, *Projections of Demand for and Costs of Social Care for Older People in England, 2010 to 2030, under Current and Alternative Funding Systems: Report to the Commission on Funding of Care and Support*, PSSRU Discussion Paper No. 2811, London School of Economics, London, 2011.

Zaidi, A. and T. Burchardt, "Comparing incomes when needs differ: Equivalization for the extra costs of disability in the UK," *Review of Income and Wealth*, 51(1), 89-114, 2005.

Zantomio, F., "Older people's participation in extra-cost disability benefits," *Journal of Health Economics,* 32, 320-330, 2013.

Appendix: Additional tables

Table A1: Summary statistics for standard-of-living and disability

*Standard-of-living**		
Do you (and your family/and your partner) have...	Cannot afford to...	Do not want/ have / or cannot afford to ...
Enough money to keep your home in a decent state of decoration?	0.083 (0.276)	0.101 (0.302)
Hobby or leisure activity?	0.036 (0.187)	0.254 (0.435)
Holidays away from home one week a year?	0.162 (0.368)	0.436 (0.496)
Household contents insurance?	0.049 (0.217)	0.109 (0.312)
Friends/family round for drink or meal at least once a month?	0.068 (0.252)	0.413 (0.492)
Make savings of £10 a month or more?	0.214 (0.410)	0.404 (0.491)
Two pairs of all-weather shoes for each person in the household?	0.022 (0.146)	0.038 (0.190)
Replace any worn out furniture?	0.153 (0.360)	0.323 (0.468)
Replace or repair broken electrical goods?	0.104 (0.306)	0.179 (0.383)
Money to spend each week on yourself, not on your family?	0.079 (0.270)	0.118 (0.322)
Disability		
Does this health problem(s) or disability(ies) mean that you have significant difficulties with any of these areas of your life? Please read out the numbers from the card next to the ones which apply to you.	Mean	Standard Deviation
difficulty in mobility (moving about)	0.327	0.469
difficulty with lifting, carrying or moving objects	0.301	0.459
difficulty with manual dexterity using hands for daily tasks	0.120	0.325
difficulty - continence (bladder/bowel control)	0.071	0.256
difficulty with communication (speech, hearing or eyesight)	0.089	0.285
difficulty with memory/concentration/learning/understanding	0.063	0.242
difficulty with recognising when in physical danger	0.013	0.114
difficulty with your physical co-ordination	0.109	0.312
difficulty in other area of life	0.123	0.328

Notes: Statistics computed over a sample of 8,183 FRS 2007-8 respondents. * Standard deviation in brackets.

Table A2: Sample means and standard deviations of covariates

	Mean	Standard Deviation
Age of adult last birthday	73.54	7.431
Female	59.2%	0.491
No. of years in FT education beyond school living age	1.02	1.676
Whether Partnered	58.0%	0.494
Area of Residence:		
North East	4.6%	0.209
North West and Merseyside	10.9%	0.312
Yorks and Humberside	8.4%	0.277
East Midlands	7.5%	0.263
West Midlands	8.3%	0.276
Eastern	8.8%	0.283
London	6.9%	0.253
South East	12.2%	0.327
South West	8.7%	0.282
Wales	5.3%	0.225
Scotland	18.4%	0.387
Net household income including disability benefits but after deducting housing costs (£ pw)[a]	283.70	176.69
Net household income excluding disability benefits and after deducting housing costs (£ pw) ('Pre-disability benefit income')[b]	*268.20*	*177.67*
Home Ownership	75.1%	0.432
Financial wealth (in £)	20,886	63,376

Notes: Sample means computed over the 8,183 respondents. Monetary values are in 2007 prices and rounded to the nearest 10p.

a. Computed as the sum across all household members of: cash income from private and state pensions, investments and savings, other market income, disability and means-tested state benefits minus income tax and housing costs. Housing costs are the sum of gross rents, council tax payments, costs of insurance on structure of property and mortgage interest payments net of housing benefit and council tax benefits.

b. Pre-disability benefits income is computed by deducting disability benefits currently received by all household members from household income.

Table A3: Standard of living and disability measurement equations

Indicator	Factor Loading	R^2
Standard of Living		
enough money to keep your home in a decent state of decoration	1.229***	*0.710*
hobby or leisure activity	0.860***	*0.545*
holidays away from home one week a year	1.139***	*0.677*
household contents insurance	0.864***	*0.547*
friends/family round for drink or meal at least once a month	0.972***	*0.604*
make savings of £10 a month or more	1.001***	*0.618*
two pairs of all weather shoes for each person in the HH	0.895***	*0.564*
replace any worn out furniture	1.789***	*0.838*
replace or repair broken electrical goods such as fridge, washing machine	1.615***	*0.809*
money to spend each week on yourself, not on your family	1.080***	*0.654*
Disability		
difficulty in mobility (moving about)	2.138***	*0.840*
difficulty with lifting, carrying or moving objects	2.435***	*0.872*
difficulty with manual dexterity using hands for daily tasks	1.327***	*0.669*
difficulty - continence (bladder/bowel control)	0.766***	*0.402*
difficulty with communication (speech, hearing or eyesight)	0.656***	*0.330*
difficulty with memory/concentration/learning/understanding	0.813***	*0.431*
difficulty with recognising when in physical danger	0.737***	*0.384*
difficulty with your physical co-ordination	1.382***	*0.686*
difficulty in other area of life	0.465***	*0.198*

Significance: * = 10%; ** = 5%; *** = 1%; R^2 is the squared correlation between the indicator (S_q or D_k) and the latent variable (φ or η). Estimates are from the quadratic in ln(Y) model specification.

Table A4: Estimates of the structural parameters of the disability equation

Covariate(s):	Coeff.	S.E.
Spline age 73†	0.033***	*0.002*
Spline age 73 and over †	0.033***	*0.003*
Female	-0.005	*0.028*
Post-compulsory schooling	-0.036***	*0.009*
(ln) pre-disability benefit income	-0.114***	*0.028*
Home ownership	-0.299***	*0.034*
(ln) financial wealth	-0.029***	*0.004*

Note: Significance: * = 10%; ** = 5%, *** = 1%; †Cut-off set to 73, the median age in the sample. Model also includes controls for region of residence and marital status. R^2=0.127. Estimates are obtained using the quadratic in ln(Y) model specification.

Table A5: Parameter Estimates from the Standard of Living Equation in the Three Variants

	Model (1) Linear in Y		Model (2) Linear in ln(Y)		Model (3) Quadratic in ln(Y)	
Covariate(s):	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
η	-0.277***	*0.020*	-0.29***	*0.020*	-0.281***	*0.020*
(in s.d. units)	*-0.233****	*0.016*	*-0.254****	*0.016*	*-0.236****	*0.016*
linear in Y	0.003***	*0.000*				
(in s.d. units)	*0.003****	*0.000*				
linear in ln(Y)			0.790***	*0.036*	-3.424***	*0.272*
(in s.d. units)			*0.631****	*0.026*	*-2.610****	*0.201*
quadratic in ln(Y)					0.400***	*0.027*
(in s.d. units)					*0.307****	*0.019*
Female	-0.005	*0.033*	0.002	*0.033*	-0.004	*0.033*
(in s.d. units)	*-0.004*	*0.026*	*0.002*	*0.027*	*-0.003*	*0.026*
Spline age 73[a]	0.034***	*0.004*	0.03***	*0.004*	0.034***	*0.005*
(in s.d. units)	*0.026****	*0.003*	*0.025****	*0.004*	*0.027****	*0.004*
Spline age 73plus[a]	0.037***	*0.004*	0.038***	*0.004*	0.037***	*0.004*
(in s.d. units)	*0.029****	*0.003*	*0.031****	*0.003*	*0.029****	*0.003*
Post-compulsory schooling	0.043***	*0.012*	0.062***	*0.011*	0.041***	*0.012*
(in s.d. units)	*0.034****	*0.009*	*0.051****	*0.009*	*0.032****	*0.009*
Home owner	0.497***	*0.037*	0.481***	*0.037*	0.486***	*0.037*
(in s.d. units)	*0.39****	*0.030*	*0.393****	*0.030*	*0.382****	*0.029*
Married/cohabiting	-0.195***	*0.041*	-0.222***	*0.039*	-0.238***	*0.040*
(in s.d. units)	*-0.153****	*0.031*	*-0.181****	*0.032*	*-0.187****	*0.031*

Notes: Significance: * = 10%; ** = 5%, *** = 1%. a. Cut-off set to 73, the median age in the sample. All models also contain controls for region of residence.

Figure A1: Estimated form of the Income-SoL profile

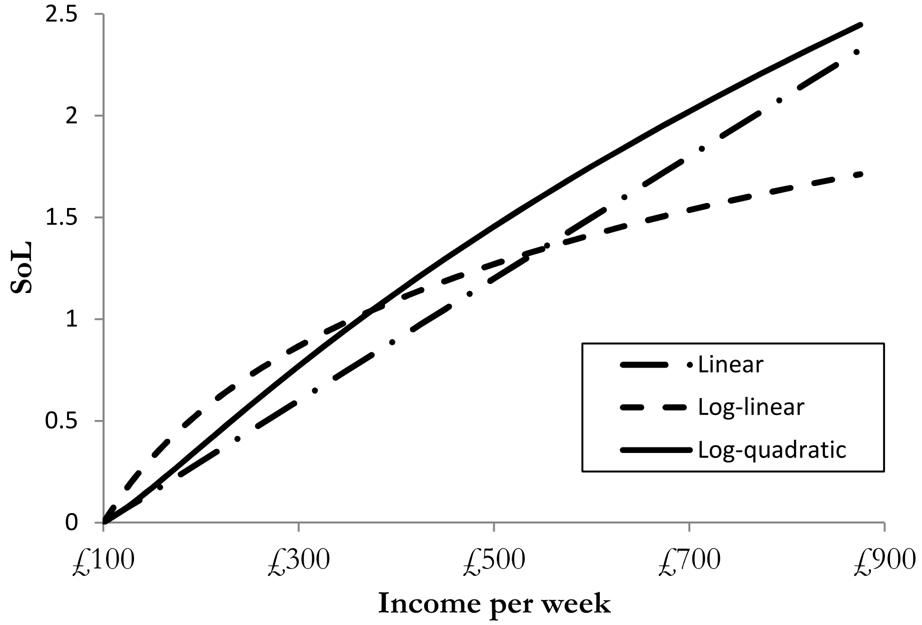
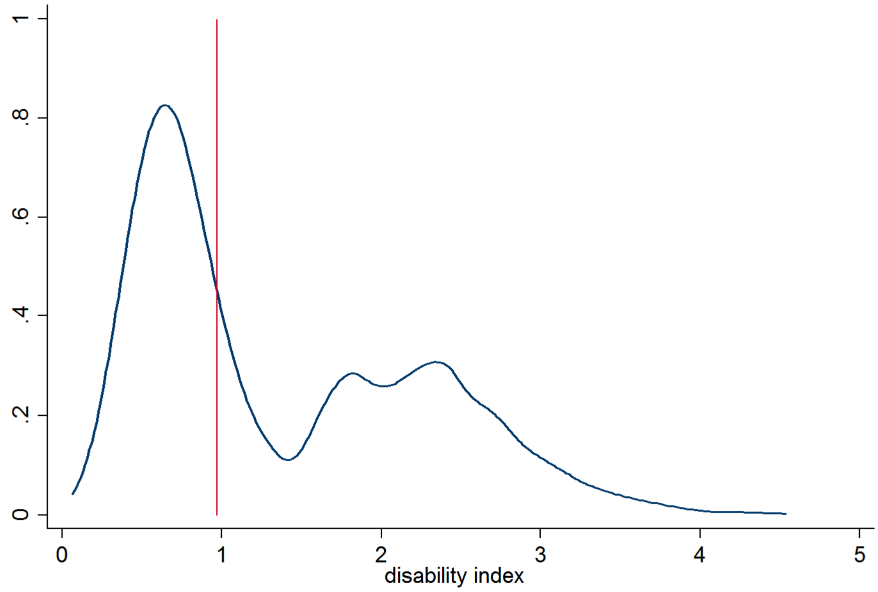


Figure A2: Kernel Density Estimator of η



Notes: vertical line represents the reference level of disability (D_0).

Chapter 4:

Socio-economic disparities in cohort-year trends in disability and receipt of disability benefits at old-age: Evidence from the UK*

Abstract: Public programmes of support for disabled people are facing increasing financial pressure in contemporary societies. A fundamental question is how much of their growth can be explained by trends in the underlying prevalence and severity of disability. A two-latent factor structural equation approach is employed to estimate the (birth-)cohort-year effects in physical and cognitive functionings and in the receipt of non means-tested cash disability benefits (DBs) for older people born between 1924 and 1945. We found that the overall slightly increasing cohort-year trend in physical and cognitive disability hides diverging trends by socio-economic status (SES), with relevant indirect effects on DBs receipt. The direct cohort-year effect on DBs receipt is mainly attributable to an increase among the better educated individuals. Results have important implications for current and planned policy reforms aimed at supporting people with care needs.

Keywords: disability, disability benefits, birth-cohort trends, latent factor, structural equation model.

* I am grateful for comments and suggestions from Ruth Hancock and Stephen Pudney. This research was supported by the Economic and Social Research Council (grant no. ES/K003852/1: "Disability and care needs in the older population: disability benefits, social care and well-being"). Data from the Family Resources Survey (FRS) are made available via the Department of Work and Pensions through the UK Data Archive. Material from the FRS is Crown Copyright and has been used by permission. The responsibility for the analysis and the views expressed lies entirely with the author.

1 Introduction

Like the UK, many countries have experienced substantial growth in the number of people receiving disability-related benefits. Health/disability status determines eligibility for disability programmes, but it is still debated how much of the growth in their rolls is down to demographic and epidemiological pressures. The general conclusion, at least for insurance-based disability programmes[49] targeted to the working age population, is that the growth in their rolls has primarily been driven by factors unrelated to the ageing of the population and its underlying disability (Autor, 2011; Black *et al.*, 2002; Bound & Burkhauser, 1999; Burkhauser & Daly, 2011; Burkhauser & Daly, 2012; Burkhauser *et al.*, 2014; Haveman & Wolfe, 1984; Juhn, 1992; McVicar, 2008; Parsons, 1980), though a recent analysis of US data has challenged this finding (Pattison & Waldron, 2013).

Much less attention has been paid to the dynamics of social security benefits that are intended to partially compensate for the extra costs of living with a disability. In the UK, there are two alternative non-contributory tax-free non-means-tested social security disability-related cash benefits (DBs): the *Disability Living Allowance* (DLA), claimable for disabled people aged 16 to 64 (although

[49] Insurance-based disability benefits represent earnings replacement for working age individuals who have lost their ability to work and are generally conditional on workers past earnings and contributions.

receipt can continue beyond 65), and the *Attendance Allowance* (AA), a benefit similar to the DLA but claimed only from age 65.

AA and DLA are at the core of a considerable policy debate. The rapid growth in DLA rolls and its advocated *leakage* (i.e. benefits being received by the non-disabled) led to its replacement by the Personal Independence Payment (PIP), which (from 2013) has introduced a regular reassessment of the disability condition over time. AA reform prospects include the options of freezing its value (not uprated with inflation), making the benefit subject to income tax, tightening its eligibility criteria or introducing reassessment or means-testing for new claimants (Commission on the Future of Health and Social Care in England, 2014; Lloyd, 2014; Wanless, 2006). Moreover, the devolution process will provide complete autonomy to the Scottish Parliament in determining the structure and value of AA and DLA/PIP benefits.[50]

It has been reported that the bulk of DBs recipients in the UK are older people (Berthoud, 2009; Burchardt, 1999; Falkingham *et al.*, 2010; HMSO, 1988). During the period 2002 to 2012, the number of older recipients increased by more than half a million (+31%), albeit no structural reforms were implemented. Population trends could explain only part of this growth: expressed as a share of the 65+ population, AA/DLA caseloads have increased by 3 percentage points (+14%).

[50] see www.smith-commission.scot.

This paper aims to contribute to a better understanding of the major driver mechanisms behind such a growth.

The natural starting point for this enquiry is the observation of trends in old-age disabilities. It is commonly assumed that advances in medicine, technology and access to public health together with increased safety at work and a lower proportion of the workforce in manual jobs mean no increase in prevalence of disability. However, the observed prevalence of disability can increase if the life-expectancy of disabled people increases, even if the onset of disability remains stable (Crimmins *et al.*, 2009; Jarvis & Tinker, 1999). It is also possible, for example, that unfavourable conditions during infancy and childhood for the oldest cohorts had preselected the strongest members, thus suggesting an increasing trend. Additionally, it has been argued that increasing exposure to unhealthy environments (e.g. the obesogenic features of the modern environment) and associated conditions in younger compared with older cohorts, might have serious implications for age at onset of chronic health problems and therefore in the prevalence of functional disabilities in later life (Crimmins & Beltrán-Sánchez, 2011; Martin *et al.*, 2010; WHO, 2011). Not surprisingly, empirical evidences on trends in disability is mixed, with no consensus yet emerging.

The prevalence of disability is known to vary between subgroups of the population. Regardless of the socio-economic status (SES) indicator in use, results tend to show that older people with low SES are more likely to experience health problems and to die at a younger age than their counterparts (for a review see e.g. Feinstein (1993);WHO (2014)). However, the strong relation between SES and health (the so-called SES gradient in health; Deaton (2002)) is clouded when considered in the context of the ageing population; some prior studies report a convergence of health inequality by SES, while others demonstrate a persistent or diverging gap over time (e.g., Freedman *et al.*, 2002; Martin *et al.*, 2012; Morciano *et al.*, 2015; Schoeni *et al.*, 2001; Schoeni *et al.*, 2005).

The key interest for our analysis is whether there is any trend in SES inequalities in disability that might lead to relevant trends in the composition of those potentially entitled to DBs. This is important mainly for two reasons. First, despite the non-means-tested nature of DBs, it has been found that their take-up is higher among older people with low SES than those with more education or higher incomes and wealth (Hancock *et al.*, 2015; Hancock & Pudney, 2014; Morciano *et al.*, 2015; Pudney, 2010). Therefore the impact on DB rolls would be rather different in magnitude if an epidemiological pressure came from low SES individuals, rather than from those better off. Secondly, given the significant changes in the living conditions of the oldest population, it is also important to assess whether, conditional on disability and other relevant characteristics, inequalities in DBs take-up (due, for example, to stigma effects) are widening or not. In terms of "target efficiency" of the system, the answers to such questions have profoundly different policy implications.

By following multiple birth-cohorts over multiple time points, we examine (birth-) cohort-year effects in physical and cognitive functioning and in the receipt of DBs for the 65 and over population in the UK. In doing so, we use individual-level data from those born between 1924 and 1945 as observed in the Family Resource Survey (FRS) carried out from 2002/3 to 2011/12. These data contain detailed information of relevance for our study and have the advantage of covering a large national non-institutionalised population sample which yields more precise estimates of the phenomena of interest.

We employ a latent variable structural approach. It incorporates a latent variable representation of the individual's unobservable and multidimensional disability in a system of structural equations where the latent dimensions of disability (together with observable characteristics) determine receipt of DBs. The individual's disability is characterised by correlated physical and cognitive dimensions that are measured by potentially error-contaminated self-reported functional difficulty (FD) indicators.

Historical trends are driven by a confluence of age, (birth-) cohort, and time effects. Regardless of the data available, attempts to capture all three effects are faced with the identification problem that a person's age added to their birth year gives the current year, so that there is an exact linear relationship between the age, cohort, and time effects. We addressed this linear dependency by means of an additively separable age-cohort specification which assumes period effects are the same for each age level and cohort. By exploiting the range of different SES indicators in the data (measures of educational attainment, income components, and home-ownership) we also assess the presence of SES-related cohort diverging trends in physical and cognitive disabilities and DBs receipt, by means of interaction terms between birth-cohort and SES.

As we will document in section 4, we find a diverging gap in physical and cognitive disability between the socio-economically advantaged and disadvantaged in later life, with relevant indirect effects on DBs receipt. It would suggest that the growth rolls of DBs observed in the last decade among older people in the UK comes mainly from the significant epidemiological pressure exerted from successive cohorts of low SES individuals. Controlling for disability and SES, we also found evidence of a small, but significant, direct cohort effect on DBs receipt. It mainly comes from better-educated individuals who might be taking more advantage of their level of education in navigating the DB system or might have lowered the perceived stigma from claiming benefits.

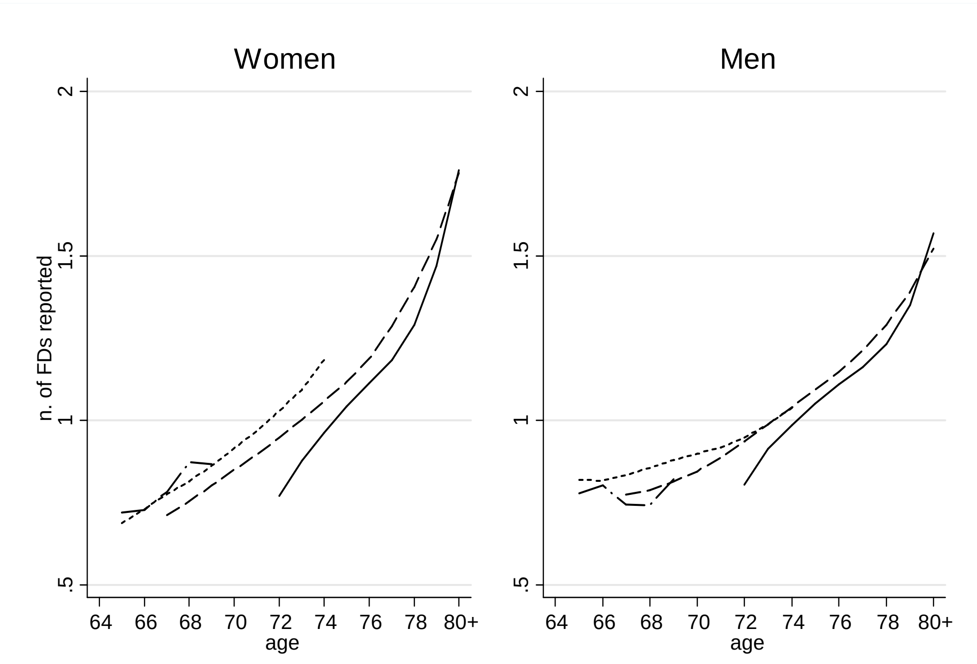
Our results have important implications for current and planned policy of the UK-DB system. The relevance of cohort-year effects also suggest that projections of the future number of disabled and the associated costs of disability programmes by shift-share analyses that use conditional rates observed at a single point in time could lead to severe underestimation.

Figure 1 illustrates the main motivation of this paper. We smooth the age trajectories of the reported number of functional difficulties (FD) (*panel a*) and of DBs receipt (*panel b*) of those cohorts observed in our data. Two facts are immediately apparent. First, different cohorts experience substantially similar profiles as they age, with both the number of reported FD and the probability of receiving DBs increasing sharply with age. Second, at a given age, successive cohorts of older people (women in particular) experience a systematic increase in the average number of FD reported as well as in the receipt of DBs. This may reflect a general tendency in the reporting propensity; alternatively it may reflect real differences between cohorts.

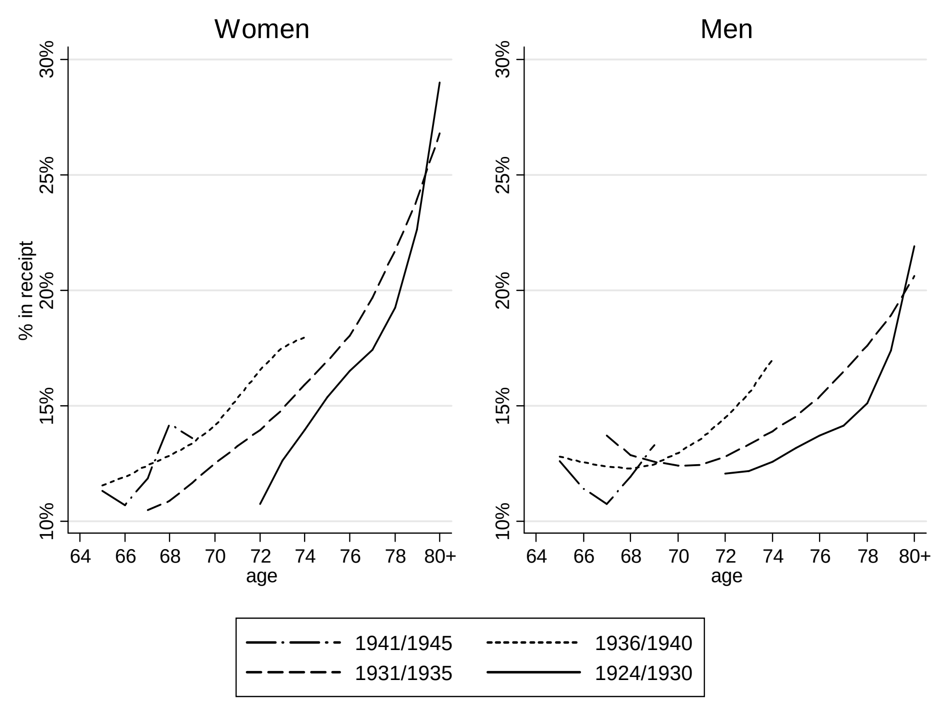
The rest of the paper is set out as follows. Section 2 describes the statistical approach in use. The pseudo-panel used for the empirical application is presented in section 3. In section 4 we report the main findings. Finally, section 5 summarises and highlights the main policy implications.

Figure 1: Age trajectories of the number of FD reported and the rate of receipt of DBs by birth-cohort and gender

Panel a)



Panel b)



Notes: Local weighted averages generated by gender and grouped-birth-cohorts specific smoothed local linear regressions (Cleveland, 1979) with bandwidth set equal to 0.8. *Source:* see section 3 for details.

2 Model specification & main assumptions

Consider a multi-equation model with $\boldsymbol{\eta}_i = (\eta_{i1} \dots \eta_{iQ})$ being an unobserved Q-dimensional "true" disability and r_i being the latent propensity to receive DB (AA or DLA) for a representative sample of individuals aged 65 and over, $i = 1, \dots, N$:

$$\eta_{iq} = \boldsymbol{\beta_{1q} X_i} + \varepsilon_{1iq} \tag{1}$$

$$r_i = \boldsymbol{\beta_2 X_i} + \gamma_1 \eta_{i1} + \dots + \gamma_Q \eta_{iQ} + \varepsilon_{2i} \tag{2}$$

where $\varepsilon_1, \varepsilon_2$ have zero mean, unit variances and are serially independent, $\boldsymbol{X}$ is a vector of exogenous variables assumed to be correlated with the outcome of interest. For identification, either at least one exogenous variable in (1) does not appear in (2) or ε_1 and ε_2 are assumed independent (Hancock *et al.*, 2015).

In this model, η_{iQ} and r_i are not directly observable. Instead we have a vector of functional disability indicators $\boldsymbol{D}$ of length J, reflecting disability $\eta_{i1} \dots \eta_{iQ}$. We assume that $\boldsymbol{D}$ contains binary indicators but the framework can be easily extended to Likert-scale response indicators without loss of generality.

The measurement component for disability can be expressed as follow:

$$D_{ij} = \mathbf{1}\left(\lambda_{j1}\eta_{i1} + \dots + \lambda_{jQ}\eta_{iQ} + \xi_{ij}\right) \tag{3}$$

where the parameters λ_{jq} are factor loadings associated with the D_{ij} indicators. The indicator function $1(.)$ maps the latent index on the right-hand side of the measurement equation (3) into the observed indicators of disability. Identification of the system of equations in (3) is achieved by imposing $E\left(\xi_j\right) = 0$ and $Cov\left(\eta_q, \xi_j\right) = 0$ and normalising the scale of the q-latent index by constraining one of its factor loading or by setting its residual variance to 1.

DB receipt is modelled by a probit specification. B_i is a binary variable indicating the actual receipt of DBs (so $B_i = 1$ if the individual receives DBs and 0 otherwise) we have that $B_i = 1$ if $r_i \geq 0, B_i = 0$ otherwise.

As detailed in section 3, the vector $\boldsymbol{X}$ includes the observable personal characteristics required to define the cohort-specific SES gradient in disability and DBs receipt.

Linear models that attempt to capture the contemporaneous effects of age (a_{it}), birth-cohort (c_i) and time (t) face an identification problem following from the identity, $a_{it} + c_i = t$.[51] Age represents biological and physiological factors associated with the ageing process, affecting equally all individuals of the same age. Year of observation (period) represents contemporaneous conditions that have an effect on individuals of all ages.[52] Year of birth (cohort) represents the *cumulative* influence of past conditions on a group of individuals that ages with a similar timing.

[51] See e.g. Glenn (1976). Bell and Jones (2013) provide a recent review of the numerous, but still unsatisfactory, attempts to "solve" this identification problem.

[52] These may be associated, for example, with advances in medical knowledge, development of new diagnostic methods, access to health facilities, improvements in sanitary conditions, life conditions and other environmental conditions.

Our analysis involves two assumptions on age, cohort, and period variables. The first assumption rules out the cross terms among these variables.[53] The second assumption is that period effects capture transient events which are assumed to be the same for each age level and cohort. Such events can be absorbed in the residual terms $\varepsilon_1, \varepsilon_2$, allowing the effects of the inter-cohort and inter-age disparities to be isolated. If period effects actually show a trend, our estimates of a_{it} and c_i needs to be reinterpreted as composite effects of a_{it}, c_i and t.[54] Since we do not want to rule out this potential source of bias, we refer to *cohort-year* and *age-year* effects throughout the paper without imposing further (untestable) assumptions.

In our application we also tested for the presence of SES-specific cohort-year trends, by adding interactions between c_i and the SES indicators used in the analysis. It is important to note that under the assumption that the transient events occur uniformly across all SES groups,[55] unobservable period effects would not bias the coefficients associated with the interaction terms.

[53] In fact, if the effects are not assumed to be additive (e.g. younger people are more/less exposed than the oldest ones to advances in medical knowledge), they would lead to a more complex set of assumptions that age effects vary systematically among the periods and cohorts, that period effects vary systematically among age levels and cohorts, and that cohort effects vary systematically among age levels and period.

[54] For example, positive period effects in disability may be driven by increased exposure to risk factors (i.e. obesity) or by medical advances that increase life expectancy of the disabled population without reducing their level of disability. Assuming positive age and birth-cohort effects, such period effects would bias upward their estimated effects. Conversely, negative period effects due for example to medical advances that prevent the onset of disability or that reduce the disabling effects of chronic conditions would bias downward the estimated effects of a_{it} and c_i.

[55] SES-differentials in the exposure to the transient events might occur because of unequal access to medical advances between the socioeconomically advantaged and disadvantaged. However, this should not be the case - at least in principle - in countries like the UK with universal public healthcare systems.

We estimate the system comprising equations (1) and (3) simultaneously allowing for the discrete nature of the dependent variables, using robust maximum likelihood.[56] The path diagram of the estimated model is available in the Appendix Figure A1. Gender-specific models were estimated to allow for possible gender differences in the reporting of FDs (Crimmins *et al.*, 2011; Oksuzyan *et al.*, 2010; Zaninotto *et al.*, 2010).

3 Data

3.1 The Family Resources Survey

We used data from a pooling of ten years of the Family Resources Survey (FRS). Despite the cross-sectional nature of the FRS data, it has the advantages of covering a large national population sample and containing a full range of questions relevant to the study. The FRS is sponsored by the Department for Work and Pensions (DWP) (Department for Work and Pensions, various years), is used to derive official income and poverty statistics (Department for Work and Pensions, 2013b) and provides the basis for most official statistics on welfare and disability programme targeting (Kasparova *et al.*, 2007). Each cross-sectional survey uses the Postcode Address File (PAF) as a sampling frame, and data are collected mainly in face-to-face interviews, performed by trained interviewers,

[56] We use the command `gsem` with robust (unclustered) estimator as implemented in *Stata 13.1 MP*. We used a probit link function in equation (3).

from a large representative sample of individuals (on average about 45,000 individuals aged 16+ per year) living in private households in the UK. The overall response rate was on average around 60 percent (Department for Work and Pensions, 2013a) and data were adjusted for possible differential non-response using weights constructed by DWP.[57]

Analyses were carried out using data for respondents aged 65 and over, born between 1924 and 1945[58], interviewed in one of the ten surveys carried out from 2002/3 to 2011/12.

Following Hancock *et al.* (2015), we include in the analysis proxy cases (participants who were not able to provide responses themselves) since they are likely to include some of the most severely disabled respondents. After deleting a few cases where relevant information was missing, a sample of 96,733 was selected.

Respondents to the FRS were asked whether they have any long-standing[59] illness, disability or infirmity. Respondents who answered "*yes*" were then asked if that means they have significant difficulties in the following areas of life:[60] mobility (moving about); lifting, carrying or moving objects; manual dexterity

[57] Weights control for differential response by demographic characteristics such as age, sex, marital status, region of residence, Council Tax Band (as a proxy for income) and housing tenure.

[58] To protect FRS respondents' confidentiality, age was top-coded at the age of 80, necessitating the exclusion of those born before 1924.

[59] Long-standing is used here to refer to "anything that has troubled you over a period of at least 12 months or that is likely to trouble you over a period of at least 12 months".

[60] The following question wording is used: "*Does this/do these health problem(s) or disability(ies) mean that you have substantial difficulties with any of these areas of your life? Please read out the numbers from the card next to the ones which apply to you.*"

(using your hands to carry out everyday tasks); continence (bladder and bowel control); memory or ability to concentrate, learn or understand; recognising when you are in physical danger; physical co-ordination (e.g. balance); other health problem or disability. All eight self-reported functional difficulty (FD) indicators available in the FRS are included in $\boldsymbol{D}$. A binary variable on whether interview was taken by proxy is also included in $\boldsymbol{D}$, possibly capturing the health-related dimensions of not taking part in the interview in person.

Observations are made on individuals of age 65 and over. The sample was divided into birth-cohorts, with some cohorts observed in more time periods than others because of the imposed age restriction. Table 1 presents a Lexis diagram for the observed 21 birth-cohorts by age and year of the interview. To simplify the exposition of results, c_i was set to 1 for the first birth-cohort in our sample, the 1924 cohort and increased by 1 for each successive year-cohort.

As indicators of SES, we used the level of education (compulsory education versus post-compulsory education),[61] home-ownership and net household income. Income was constructed as the sum of wages and salaries, self-employment income, public pensions, social security income and capital income (interest, rent,

[61] It should be noted that the distribution of educational attainment among today's older people is likely to be highly skewed. This is because the majority of them left school at the minimum permitted age. Educational attainment may therefore discriminate only between the most advantaged and the rest of the older population. We therefore decided not to differentiate further beyond the compulsory education to avoid unreliability of estimates due to small sample cell sizes.

dividends, private pensions and annuities), less income tax payments. We excluded income from disability benefits which, in the UK, are paid by the state to disabled older people in recognition of the extra costs that disability brings and so are a consequence, not a cause, of disability (Hancock *et al.*, 2015; Hancock & Pudney, 2014). Home-ownership and income from capital represent returns on assets accumulated over the lifecycle and are consequently good indicators of past access to resources with an expected cumulative positive influence on health (Morciano *et al.*, 2015). Income is aggregated across all household members and divided by the square root of the number of people in the household to adjust for differences in household size.[62] Given the skew of the income distribution, we follow common practice and enter income in log transformed form. We also control for country (within the UK) of residence of the respondents.

[62] Since the majority of households in our analysis consist of one or two adults, applying other commonly used scales, such as the OECD modified equivalence scale, would not yield substantially different results.

Table 1: Lexis diagram of the observed cohorts by age and year of interview

Cohort	Age														
of birth	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80+
1924														2002	2003	2004
1925													2002	2003	2004	2005
1926												2002	2003	2004	2005	2006
1927											2002	2003	2004	2005	2006	2007
1928										2002	2003	2004	2005	2006	2007	2008
1929									2002	2003	2004	2005	2006	2007	2008	2009
1930								2002	2003	2004	2005	2006	2007	2008	2009	2010
1931							2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
1932						2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
1933					2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	
1934				2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012		
1935			2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012			
1936		2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012				
1937	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012					
1938	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012						
1939	2004	2005	2006	2007	2008	2009	2010	2011	2012							
1940	2005	2006	2007	2008	2009	2010	2011	2012								
1941	2006	2007	2008	2009	2010	2011	2012									
1942	2007	2008	2009	2010	2011	2012										
1943	2008	2009	2010	2011	2012											
1944	2009	2010	2011	2012												
1945	2010	2011	2012													

Source: Data on 65+ respondents born between 1924-1945, interviewed in the FRS survey from 2002/3 - 2011/12.

3.2 Descriptive statistics

Table 2 shows the main characteristics of the study population disaggregated by gender. Differences between genders were almost all significant at the 1% level. Women, who represented about 55% of the total sample, reported higher disability prevalence than men at both the extensive (the probability of being functionally disabled) and intensive margins (the severity of functional disability, among disabled). They also reported higher prevalence of the four most common types of FD (mobility, lifting, dexterity, co-ordination), although three less common types (incontinence, communication, memory) were reported a little more frequently by men. Specifically, women reported higher prevalence of 1+ and 4+ FDs and, among disabled, higher average severity (number of FDs reported) compared with men, reflecting significant gender differences for all FD indicators with the exception of the item "recognising when in physical danger", for which no significant difference was found.

The median age of respondents in the sample was 73 for men and 74 for women. Mean household income (expressed in 2012 prices) was £367 per week for men and £321 for women. The majority of respondents were home-owners (80% for men; 76% for women), most had a post-compulsory school qualification (67% for men; 65% for women), and most were resident in England (84%).

Socio-economic differentials in the prevalence of FDs were marked as shown in Table 3. The proportions reporting at least one FD, four or more FDs and the average number of reported FDs among those with at least one FD (severity), were higher among people without post-compulsory education, non-home-owners and those in the poorest quartile of the income distribution. Receipt of DBs shows an even great gradient with SES, reflecting the higher prevalence of functional disabilities or the stronger incentive to claim DBs for low-income households.

Table 4 reports the prevalence and severity of functional disability, and means of the SES variables according to birth-cohort and age group. For each age group apart from 80+, the prevalence of disability was slightly lower in successive birth-cohorts. However, the severity of disability among those who reported it, increased significantly for successive cohorts within all age groups. Successive birth-cohorts of older people reported significant improvements in SES, mainly in the percentage of individuals reporting post-compulsory education.

Table 2: Functional Difficulties (FDs) and selected socio-economic indicators in the pooled sample of FRS

	Men		Women		
	Mean	Standard error	Mean	Standard error	Difference
Mobility	31.2%	0.463	35.7%	0.479	-0.034***
Lifting	28.3%	0.450	33.0%	0.470	-0.036***
Dexterity	10.9%	0.311	14.6%	0.353	-0.034***
Co-ordination	9.9%	0.299	11.5%	0.319	-0.011***
Communication	9.8%	0.297	8.8%	0.283	0.014***
Incontinence	8.4%	0.277	7.5%	0.263	0.011***
Memory	7.8%	0.268	7.0%	0.255	0.011***
Recognise when in danger	1.6%	0.126	1.9%	0.137	-0.001
No FDs reported	*56.7%*	*0.495*	*53.9%*	*0.499*	*0.019****
1 or more FDs reported	*43.3%*	*0.495*	*46.1%*	*0.499*	*-0.019****
4 or more FDs reported	*9.4%*	*0.292*	*10.8%*	*0.310*	*-0.008****
number of FDs (among disabled)	*2.49*	*1.516*	*2.60*	*1.516*	*-0.073****
Median age [a]	73	5.114	74	5.246	-1***
Equivalised pre-disability benefit household income[b]	366.72	322.57	321.18	272.07	41.122***
Post-compulsory school	67.9%	0.467	65.0%	0.477	0.008**
Home-ownership	79.9%	0.401	75.7%	0.429	0.04***
England	83.9%	0.368	83.3%	0.373	0.014***
Wales	5.5%	0.227	5.4%	0.225	0.002
Scotland	8.2%	0.275	8.8%	0.283	-0.013***
Northern Ireland	2.4%	0.154	2.5%	0.157	-0.003*
In receipt of AA	7.35%	0.261	11.62%	0.320	-0.043***
In receipt of DLA	8.02%	0.272	7.68%	0.266	-0.034**
In receipt of AA or DLA	15.05%	0.358	18.95%	0.392	0.039***

Source: Weighted data on 65+ respondents born between 1924-1945, interviewed in the FRS survey from 2002/3-2011/12. *Unweighted sample size:* 52,229 women and 44,504 men. *Notes:* [a] To protect confidentiality, FRS data were released with a top-coding at the age of 80. Therefore, we report median rather than mean values. Consequently, a Pearson chi-squared test of the equality of the medians of the difference between men and women was performed. [b] (£ pw, 2012 prices) For definition of household income see text. *Level of significance:* * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 3: Prevalence and severity of disability by SES

SES indicator	Reporting at least 1 FD	Reporting at least 4 FDs	Average number of reported FDs (among disabled)	In receipt of disability benefits (DBs)
Education				
Compulsory education	56.3%	14.6%	2.70	24.6%
Post-compulsory education	39.0%	8.0%	2.45	13.5%
Home-ownership				
Non-home-owner	59.5%	15.3%	2.70	29.1%
Home-owner	40.6%	8.7%	2.49	13.8%
Quantiles of pre-disability income[a]				
Poorest 25%	49.2%	11.0%	2.55	19.7%
Richest 25%	32.4%	6.8%	2.44	8.6%
Overall	44.9%	10.2%	2.55	17.2%

Source: Weighted data on 65+ respondents born between 1924-1945, interviewed in the FRS survey from 2002/3-2011/12. *Unweighted sample size:* 52,229 women and 44,504 men. *Notes:* Differences between groups were all statistically significant at 1% level. [a] For definition of household income see text.

Table 4: Birth-cohort trends in prevalence and severity of disability and SES by age-group

Cohort of birth	Age group																			
	65-69					70-74					75-79					80+				
	Functional disability		SES indicator			Functional disability		SES indicator			Functional disability		SES indicator			Functional disability		SES indicator		
	Prevalence (a)	Severity (b)	Education (c)	Income (d)	Home-ownership (%)	Prevalence (a)	Severity (b)	Education (c)	Income (d)	Home-ownership (%)	Prevalence (a)	Severity (b)	Education (c)	Income (d)	Home-ownership (%)	Prevalence (a)	Severity (b)	Education (c)	Income (d)	Home-ownership (%)
1924	.	.	.	.	.	.	.	.	.	.	0.55	2.23	0.35	283.46	0.70	0.60	2.79	0.37	286.46	0.67
1925	.	.	.	.	.	.	.	.	.	.	0.55	2.28	0.38	285.42	0.70	0.59	2.87	0.36	283.31	0.68
1926	.	.	.	.	.	.	.	.	.	.	0.50	2.32	0.39	301.22	0.73	0.60	2.96	0.39	291.07	0.70
1927	.	.	.	.	.	.	.	.	.	.	0.48	2.34	0.37	295.69	0.74	0.61	2.91	0.41	299.16	0.72
1928	.	.	.	.	.	0.44	2.22	0.36	286.53	0.77	0.47	2.37	0.40	305.15	0.75	0.61	2.87	0.39	308.43	0.74
1929	.	.	.	.	.	0.49	2.15	0.37	311.40	0.73	0.47	2.50	0.42	318.02	0.77	0.62	2.98	0.40	309.40	0.74
1930	.	.	.	.	.	0.40	2.13	0.39	318.15	0.77	0.48	2.56	0.43	331.82	0.78	0.62	2.84	0.42	313.55	0.75
1931	.	.	.	.	.	0.41	2.28	0.45	323.76	0.79	0.48	2.47	0.44	334.77	0.78	0.60	2.93	0.43	299.07	0.76
1932	.	.	.	.	.	0.43	2.14	0.44	319.71	0.78	0.48	2.58	0.48	335.58	0.80	0.63	3.01	0.45	329.87	0.78
1933	0.39	1.94	0.60	353.00	0.78	0.39	2.30	0.59	338.21	0.79	0.46	2.64	0.60	340.80	0.79	.	.	.	.	.
1934	0.36	2.03	0.83	328.56	0.79	0.41	2.44	0.85	333.79	0.80	0.45	2.70	0.87	351.90	0.81	.	.	.	.	.
1935	0.37	2.14	0.88	340.19	0.79	0.40	2.38	0.88	351.87	0.80	0.41	2.49	0.88	335.13	0.79	.	.	.	.	.
1936	0.35	2.30	0.91	354.11	0.80	0.40	2.44	0.90	353.13	0.80	0.42	2.42	0.88	344.85	0.83	.	.	.	.	.
1937	0.36	2.19	0.92	367.81	0.81	0.40	2.57	0.91	347.47	0.81	0.42	2.62	0.92	351.38	0.82	.	.	.	.	.
1938	0.36	2.23	0.93	393.31	0.81	0.40	2.48	0.91	356.63	0.81	.	.	.	.	.	.	.	.	.	.
1939	0.35	2.39	0.93	381.99	0.81	0.38	2.46	0.93	362.13	0.81	.	.	.	.	.	.	.	.	.	.
1940	0.34	2.38	0.94	386.78	0.81	0.37	2.33	0.94	353.50	0.79	.	.	.	.	.	.	.	.	.	.
1941	0.33	2.46	0.94	395.87	0.80	0.35	2.21	0.95	358.50	0.79	.	.	.	.	.	.	.	.	.	.
1942	0.32	2.45	0.95	413.92	0.82	0.34	2.55	0.97	411.81	0.84	.	.	.	.	.	.	.	.	.	.
1943	0.32	2.45	0.95	406.14	0.81	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
1944	0.29	2.48	0.96	441.57	0.81	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
1945	0.31	2.45	0.96	421.56	0.81	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Tests for Stationarity (p-values)																				
ADF	0.99	0.16	0.62	0.75	0.04	0.99	0.53	0.52	0.89	0.59	0.81	0.48	0.97	0.66	0.27	0.65	0.00	0.99	0.01	0.86
PP	0.53	0.03	0.09	0.88	0.03	0.72	0.53	0.71	0.93	0.79	0.53	0.47	0.97	0.65	0.52	0.66	0.01	0.99	0.95	0.88

Source: Weighted data on 65+ respondents born between 1924-1945, interviewed in the FRS survey from 2002/3-2011/12. *Unweighted sample size:* 52,229 women and 44,504 men. *Notes:* [a] % of people reporting at least one FD; [b] number of FDs reported among those who reported at least one FD; [c] % of individuals reporting post-compulsory school; [d] equivalised pre-disability benefit household income (£ pw, 2012 prices). See text for the income definition. We tests for time-trends in the data using both the Augmented Dickey–Fuller (ADF) and the Phillips–Perron (PP) tests (null hypothesis of a unit root) with two lagged difference terms included in the covariate lists. Experiments with fewer or more lags in the augmented regression yield similar conclusions.

3.3 The number of factors

Disability is a multidimensional concept, with some controversy over the number of dimensions needed for empirical analysis (see e.g. Fitzgerald *et al.* (1993); Johnson and Wolinsky (1993); Spector and Fleishman (1998)) and the way that multiple domains of functioning develop together and interrelate (see e.g. Bruce *et al.* (1994); Verbrugge and Jette (1994)).

Table 5: Measures of model fit and residual variances for the one and two- factor models

Model	N. of parameters	Chi-Square χ^2	df	P-value	RMSEA	SRMSR
1-factor	9	8218.224	27	0.000	0.056	0.094
2-factor	17	574.423	19	0.000	0.017	0.023
Model compared						
1-factor against 2-factor		5478.891	8	0.000		

Notes: df (Degree of freedom); RMSEA (Root Mean Square Error Of Approximation); SRMSR (Standardized Root Mean Square Residual).

We ran an Exploratory Factor Analysis[63] for the dimensionality of disability (i.e. to extract the set of Q-latent constructs from the observed $\boldsymbol{D}$ indicators in our sample)[64] before estimating the SEM model comprising equations (1)-(3). The eigenvalues-greater-than-1 rule suggested the exploration of a two-factor model (eigenvalues 5.079 and 1.271). Table 5 shows goodness-of-fit indices and strongly

[63] In the robustness checks of an empirical investigation of target efficiency of Attendance Allowance in England, Hancock *et al.* (2015) explored the use of a two-factor measurement model to distinguish between the physical and cognitive aspects of disability. In this paper we extend that research.

[64] The analyses were run in *Mplus 7.11* using Maximum Likelihood extraction method and a geomin rotation. For identification, latent residual variances were constrained to 1.

suggests a model with two distinct disability dimensions, although the one-factor model also meets conventional standards for good fit.

Conceptually, the two-factor model provided a plausible factor structure (Table 6). The highest loadings for the first factor are for those indicators that describe mainly physical functional difficulties associated with "mobility", "lifting" and "dexterity". With the exception of the indicators "incontinence" and "co-ordination" the loadings associated with the first factor are double in magnitude than those associated with the second factor. We call this factor *physical disability* to emphasise its high correlation with the physical function indicators collected in the FRS.

Table 6: Loadings for the two factor model

	Geomin rotated loadings	
	1	2
	Physical	Cognitive
Mobility	**0.879**	0.015
Lifting	**1.036**	-0.085
Dexterity	**0.768**	0.091
Incontinence	**0.340**	**0.394**
Communication	0.146	**0.623**
Memory	0.126	**0.737**
Recognise when in danger	-0.007	**0.874**
Co-ordination	**0.512**	0.363
Interviewed by proxy	-0.274	**0.526**
Correlations among factors	0.615	

Notes: All loadings and the correlation among factors were significant at 5% level. For identification purposes, latent residual variances were constrained to 1.

The highest loadings for the second factor are for indicators of cognitive difficulties such as "recognise when in danger", "memory" and "communication". In the vector $\boldsymbol{D}$, we also included a binary indicator of whether survey information was collected by proxy for participants who were not able to provide responses themselves. We found that this indicator is strongly related to the second factor, possibly capturing the fact that, while physically impaired individuals might be more likely to be at home during the interview (Stoop, 2005), poor cognitive functioning is an obstacle to survey response. We label this second factor *cognitive disability*.

We found evidence of a weak (item loadings just above 0.30) cross-loading for "incontinence", perhaps reflecting that incontinence could be driven by both physical and cognitive impairments. Considering also its relatively small loadings, we decided to discard this indicator from subsequent analyses.[65] Recalculation of the loadings and the correlation, while not leading to substantial differences, results in a clearer factor structure[66] with all physical (cognitive) difficulty indicators loaded strongly in the physical (cognitive) factor and relatively less strongly in the other.[67]

We finally tested whether estimates of the best-fitting two-factor model are sensitive to the assumption of gender invariance. This was done by using a two-

[65] Stevens (2009), for example, recommends interpreting only factor loadings with an absolute value greater than 0.4.

[66] Allowing for a cross-loading also increases the computational complexity of the model. Previous attempts to use such a specification failed because convergence was not achieved after 216 hours by the likelihood optimiser available in STATA 13 MP16 in an e-cluster setting.

[67] A commonly employed approach to measuring disability status consists of summing the responses from $\boldsymbol{D}$. The reliability of a sum-score disability index (Cronbach's alpha) for the sample as a whole is 0.754, which lies within the bounds of what is considered to be the acceptable value, ranging from 0.70 to 0.95 (Bland & Altman, 1997; Nunnally *et al.*, 1967). The reliability of the scale constructed using the indicators which might be thought to represent physical disability is 0.760 whereas we found low reliability for the scale for cognitive disability (0.490) reflecting the complexity of its measurement.

group analysis which allows the factor loadings of the $\boldsymbol{D}$ indicators to differ according to the gender of the respondent. The goodness-of-fit tests suggested that the unrestricted model (in which the gender-specific parameters were allowed to be freely estimated) provides a better balance between model fit and parsimony.

4 Estimation results

In this section we present findings from the estimation of the gender-specific two-factor latent variable structural equation model comprising equations (1)-(3). We begin with the baseline model (*model A*) in which birth-cohort c_i was entered linearly to assess the presence of cohort-year shifts. We also include SES indicators in $\boldsymbol{X}$ to capture SES-inequality in both latent disability and receipt of DBs. We then checked the presence of SES-specific paths by birth-cohort by introducing interaction terms of c_i with SES indicators (*model B*).

4.1 Model fit

As detailed in Table 7, specification of model B was found to fit the data better for both women and men.[68]

[68] We also tested whether introducing in model B interaction terms of *age* with SES indicators provides a better fit. This model would capture the extent to which advantage protects against over-time health decline due to ageing. It would also test whether SES differential in birth-cohort trends of disability and DBs receipt remains significant when controlling for SES-age differentials. With respect to model B we got a slight improvement in AIC (271604.41 for women; 222558.86 for men) but a deterioration in the BIC criterion (272131.40 for women; 223059.11 for men) which penalises non-parsimonious models more heavily than AIC. Using this specification, the coefficients associated with c_i and SES-cohorts interactions were significant at 1% level while coefficients of age and SES-age interactions were only marginally significant. Estimates and test results are available upon requests.

Table 7: Fit measures and residual variances for the two-latent factor SEM models

Goodness of fit	Women		Men	
	Model A	Model B	Model A	Model B
Observations	52,229		44,504	
Log-likelihood	-135922	-135755	-111349	-111225
Degree of freedom (Df)	46	55	46	55
AIC	271936.54	271619.19	222790.31	222560.86
BIC	272344.26	272106.68	223190.66	223039.54

4.2 The disability model

4.2.1. The measurement equations

Estimates of the two-factor measurement model of equation (3) are shown in Table 8 for women and men separately, and for models A and B, respectively.[69]

All factor loadings λ_{jq} are positive and highly significant, with virtually no difference between the loadings estimated for models A and B. The λ_{jq} are significantly different for women and men, meaning that there are gender-specific health/disability processes or differential reporting behaviours by men and women. In all models considered, the scale of each latent disability variable was normalised by setting one of its loading to 1, leaving the variances of the latent disability variables and their correlation unconstrained. The estimated variance $\hat{\sigma}^2_{(.)}$ is found to be greater among men than women, mainly for physical disability,

[69] A sum-score index orders individuals in a very similar way to the Empirical Bayes prediction of latent disability (see later). The correlation between the two indices for physical (cognitive) disability is 0.97 (0.68) for women and 0.96 (0.71) for men with the scatterplots in Appendix Figure A2 that approximate straight lines.

η_1.[70] The covariance between the two latent factors is positive and highly significant, implying a correlation between the two dimensions of disability of about 0.33 for women and 0.26 for men.[71]

4.2.2. The structural component of disability

Tables 9 and 10 report the gender-specific coefficients $\boldsymbol{\beta_1}$ from the latent physical and cognitive disability equations.

In model A, increasing age raises the conditional mean of both latent physical η_1 and cognitive η_2 disability indices. While there is clear evidence of a negative SES-gradient with physical disability, the gradient with cognitive disability - albeit still significant - is less pronounced. There is evidence of substantially higher physical disability in Wales, but other geographical differences are modest, albeit sometimes statistically significant.

When birth-cohort is entered linearly, being born one year later is associated with an increase in both physical and cognitive disability. Cohort-year effects in physical disability are almost three times higher for women than for men. On the other hand, birth-cohort changes in cognitive disability are very similar for men and women.

---

[70] This is mainly due to the great impact that η_1 is causing to the loading "lifting" for women, being more than four times higher than the estimated impact for men. Reported prevalence of this FD among women was about 4 percentage points higher than for men (see Table 2).

[71] It should be noticed that correlation is dimensionless while covariance is expressed in units obtained by multiplying the units of the two latent variables.

Table 8: **Estimates from the measurement equations for physical and cognitive disability**

Factor Loadings (λ_j)	Women Model A	Women Model B	Men Model A	Men Model B
	Physical disability (η_1)			
Mobility	1	1	1	1
	(.)	(.)	(.)	(.)
Lifting	4.475***	4.528***	0.989***	0.995***
	(0.095)	(0.094)	(0.073)	(0.067)
Dexterity	0.694***	0.695***	0.521***	0.527***
	(0.013)	(0.013)	(0.029)	(0.027)
Co-ordination	0.521***	0.522***	0.421***	0.426***
	(0.010)	(0.010)	(0.022)	(0.020)
	Cognitive disability (η_2)			
Communication	1	1	1	1
	(.)	(.)	(.)	(.)
Memory	1.215***	1.216***	1.317***	1.317***
	(0.043)	(0.043)	(0.049)	(0.049)
Recognize when in danger	1.380***	1.384***	1.163***	1.169***
	(0.063)	(0.063)	(0.054)	(0.054)
Interviewed by proxy	0.156***	0.147***	0.116***	0.112***
	(0.014)	(0.015)	(0.014)	(0.014)
$\hat{\sigma}^2_{\eta_1}$	3.767*** (0.105)	3.740*** (0.103)	5.975*** (0.586)	5.809*** (0.516)
$\hat{\sigma}^2_{\eta_2}$	1.063*** (0.046)	1.049*** (0.046)	1.120*** (0.051)	1.111*** (0.051)
$\hat{\sigma}_{\eta_1,\eta_2}$	1.324*** (0.037)	1.307*** (0.036)	1.755*** (0.093)	1.720*** (0.085)

Significance: * = 10%; ** = 5%; *** = 1%; the first factor loadings associated with "Mobility" and "Communication" are set to 1 to normalise the scale of the latent indices η_1 (physical disability) *and* η_2 (cognitive disability). Standard errors in parenthesis.

Specification B allows us to test for the presence of SES-related birth-cohort trends. The coefficients on the interaction terms, which measure the difference in the slope of birth-cohort for different SES as compared with the slope for the reference category, indicate birth-cohort trends which differ by SES for both dimensions of disability.[72] The statistical significance of the interactions of birth-cohort and current income and home-ownership are - for both dimensions of disability - greater than those of the interactions with educational attainment, which resulted not statistically significant at conventional levels. For income, the coefficient on the interactions show that successive cohorts of high-income individuals experienced a significant reduction in latent disability level.[73] A similar result is found for the interaction of birth-cohort and home ownership, with very similar magnitudes estimated for women and men.

Figure 2 shows the kernel distribution of the Empirical Bayes (EB) predictions[74] of individuals' latent physical and cognitive disability score from model B. A similar shape is found for women and men, although the distribution for women is shifted to the right. Particularly for men, the density of the physical index is less spatially concentrated than the density of the cognitive disability index, in

[72] In linear models, the statistical significance of the interaction effect can be tested with a single t-test on the coefficient associated with the interaction (Ai & Norton, 2003).

[73] As in many other studies, the analysis relies on the reliability of self-reported disability. In the absence of objective measures of disability or anchoring vignettes (d'Uva *et al.*, 2011; King *et al.*, 2004) we are not able to investigate the possibility that SES differences in reporting disability have changed across birth-cohorts.

[74] EB predictors of the latent variables η_1 and η_2 are the means of the empirical posterior distribution with the parameter estimates $\boldsymbol{\beta}_{\mathbf{1}(.)}$ replaced with their estimated model parameters $\widehat{\boldsymbol{\beta}_{\mathbf{1}(.)}}$.

line with the greater variance estimated for η_1 than for η_2 and with the higher value of $\hat{\sigma}^2_{\eta_1}$ estimated among men than women (see Table 8).

Table 9: Estimates from the disability equations, Women

	Physical disability		Cognitive disability	
	Model A	Model B	Model A	Model B
Age	0.077***	0.075***	0.067***	0.066***
	(0.002)	(0.002)	(0.004)	(0.004)
Post-compulsory school	-0.282***	-0.297***	-0.157***	-0.202***
	(0.012)	(0.022)	(0.020)	(0.038)
Log household income[a]	-0.146***	0.178***	0.040*	0.303***
	(0.009)	(0.020)	(0.018)	(0.036)
Home ownership	-0.564***	-0.257***	-0.252***	-0.047
	(0.013)	(0.021)	(0.019)	(0.035)
Scotland	-0.041**	-0.042**	-0.032	-0.032
	(0.014)	(0.014)	(0.022)	(0.022)
Wales	0.417***	0.416***	0.027	0.026
	(0.020)	(0.020)	(0.037)	(0.037)
Northern Ireland	-0.064**	-0.075***	-0.105**	-0.110**
	(0.022)	(0.022)	(0.036)	(0.036)
Birth-cohort	0.028***	0.212***	0.017***	0.172***
	(0.002)	(0.010)	(0.003)	(0.017)
Birth-cohort * post-compulsory school		-0.002		0.003
		(0.002)		(0.004)
Birth-cohort * income		-0.029***		-0.026***
		(0.002)		(0.003)
Birth cohort * home ownership		-0.031***		-0.022***
		(0.002)		(0.003)

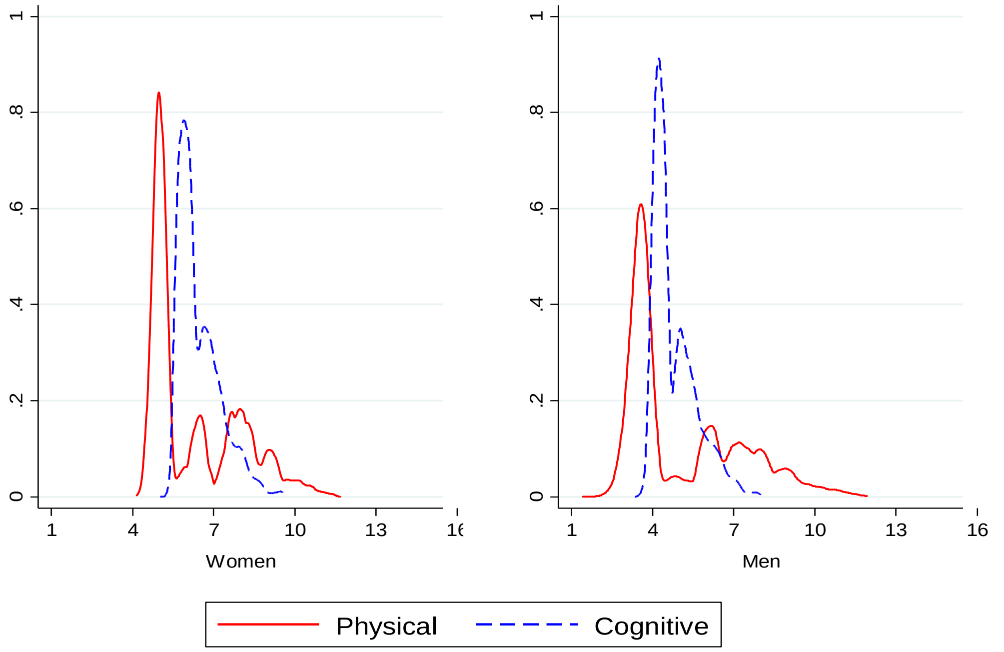
Notes: [a] For definition of household income see text. See Table 8 for estimated variances of η_1 and η_2 and their estimated correlation. Standard errors in parenthesis. *Level of significance:* * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 10: Estimates from the disability equations, Men

	Physical disability		Cognitive disability	
	Model A	Model B	Model A	Model B
Age	0.074***	0.072***	0.064***	0.064***
	(0.006)	(0.006)	(0.004)	(0.004)
Post-compulsory school	-0.254***	-0.345***	-0.049*	-0.093*
	(0.032)	(0.063)	(0.022)	(0.044)
Log household income[a]	-0.581***	0.062	-0.191***	0.006
	(0.034)	(0.053)	(0.018)	(0.038)
Home ownership	-0.850***	-0.447***	-0.241***	-0.009
	(0.050)	(0.067)	(0.021)	(0.043)
Scotland	-0.016	-0.021	-0.036	-0.038
	(0.033)	(0.033)	(0.025)	(0.025)
Wales	0.533***	0.524***	0.079*	0.077*
	(0.059)	(0.058)	(0.039)	(0.039)
Northern Ireland	-0.038	-0.059	-0.182***	-0.188***
	(0.049)	(0.050)	(0.040)	(0.041)
Birth-cohort	0.009*	0.353***	0.018***	0.130***
	(0.004)	(0.027)	(0.003)	(0.018)
Birth-cohort * post-compulsory school		0.002		0.002
		(0.006)		(0.004)
Birth-cohort * income		-0.056***		-0.017***
		(0.005)		(0.003)
Birth-cohort * home ownership		-0.036***		-0.022***
		(0.001)		(0.004)

Notes: [a] For definition of household income see text. See Table 8 for estimated variances of η_1 and η_2 and their estimated correlation. Standard errors in parenthesis. *Level of significance:* * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure 2: Kernel density distribution of the Empirical Bayes (EB) predictions of the latent disabilities scores by gender



In Figure 3, we compare the implications of the estimated model B, for illustrative hypothetical men and women, living in England at the age of 73. A set of gender-specific diagrams show the separate impacts of education (*a*), income (*b*) and home-ownership (*c*), before showing the joint impact of SES on cohort-year trends (*d*).

Graph (a) of Figure 3 reports the gender-specific estimated trend for illustrative individuals with and without post-compulsory education, all with median income and assumed to be home-owners. Three important messages clearly emerge. First, the level of physical and cognitive disability is higher for the lower-educated individual. Secondly, the educational gap can be observed at any birth-cohort and it is more apparent for physical disability (reflecting the higher estimated association of educational attainment with η_1 than with η_2). Thirdly, the upward birth-cohort trend in physical disability is almost independent of educational level (i.e. the two lines are parallel). The birth-cohort trend of η_2 is only slightly upward, at a rate which is slightly higher among those with post-compulsory education.

Graph (b) in Figure 3 isolates the effect of income. The representative individuals are those at the 25th (low), 50th (median) and 75th (high) percentiles of the income distribution observed in the whole sample.[75] All individuals are assumed to have only compulsory education and are home-owners. The between-cohort income gaps in terms of both physical and cognitive disability are increasing. The trend in the predicted mean of η_1 and η_2 across birth-cohorts is steep and upward for the low-income man and woman but less pronounced for the median-income individuals. For the high-income woman the trend is almost flat. For the high-income man it is significantly downward in η_1 and almost flat for η_2.

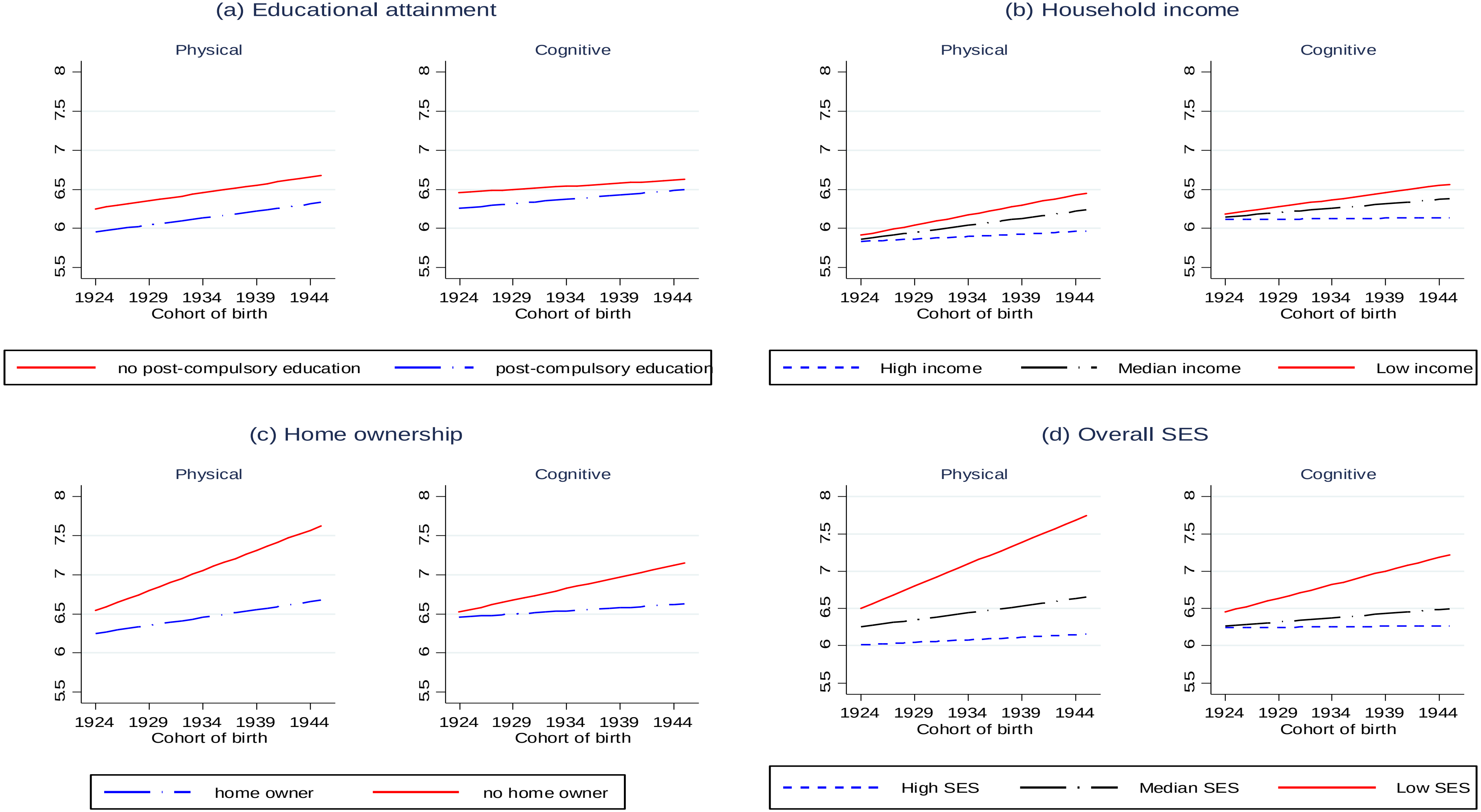
Graph (c) isolates the effect of home-ownership. The two illustrative individuals used for this exercise are assumed to have median income and have only compulsory education. The trend in the predicted mean of η_1 and η_2 across birth-cohorts for the non-home-owner man and woman is steep and upward. The homeowner benefits from a more favourable upward birth-cohort trend in both

[75] Their income (in logs) is 5.30, 5.56 and 5.95, respectively.

dimensions of disability, which becomes virtually flat for the woman with cognitive disability and the man with physical disability.

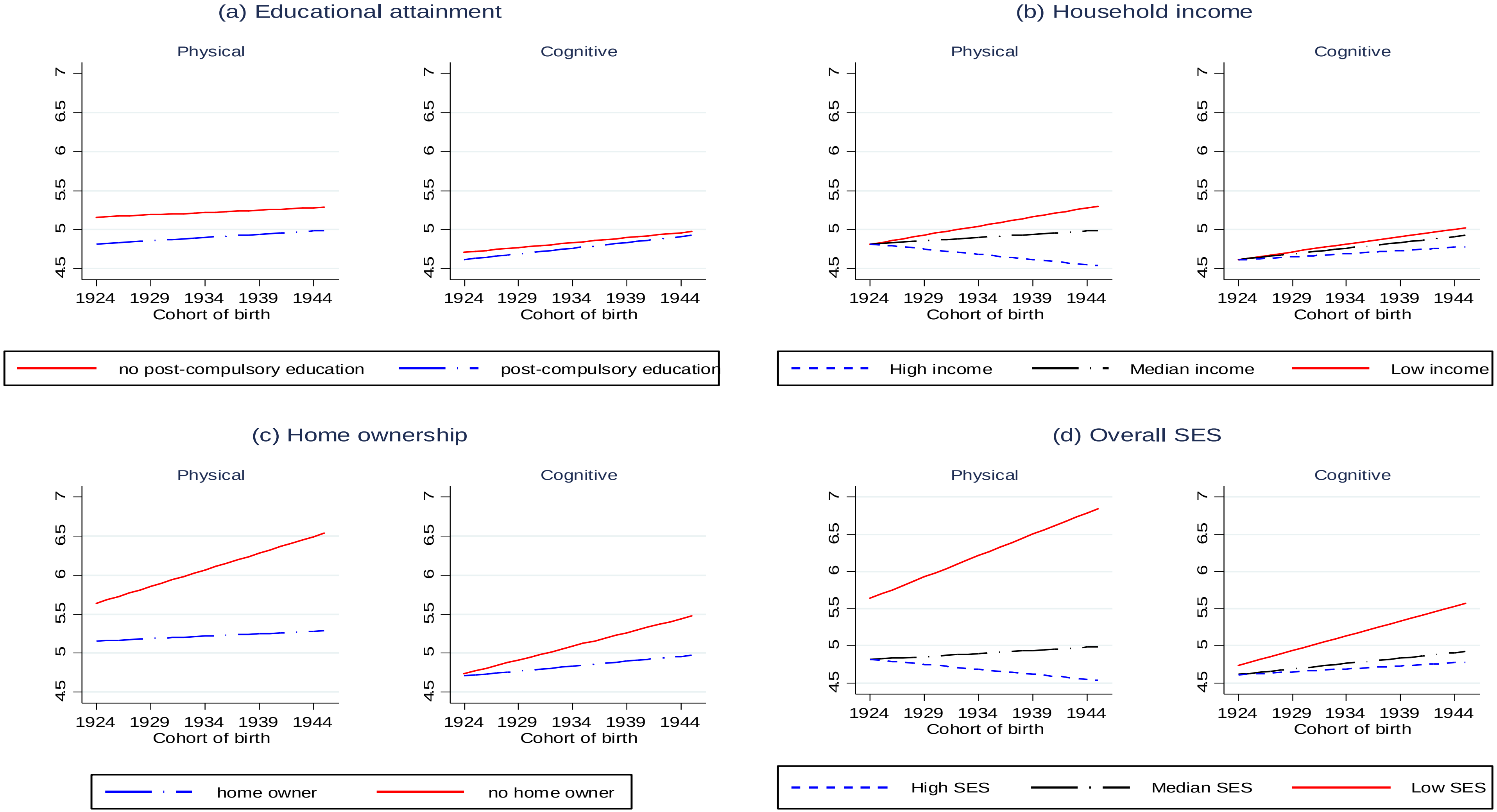
The inspection of the effect of the three dimensions of SES in this piecemeal way does not take into consideration that, in reality, such dimensions are highly correlated and therefore it might provide a partial – even if useful - representation of the underlying birth-cohort trends. The final graph (d) shows birth-cohort trends for three illustrative men and women who might appear to be more representative. As before, they are assumed to live in England at the age of 73: at the 25th (low SES), 50th (median SES) and 75th (high SES) percentiles of the income distribution in the sample. Both median- and high-SES individuals have post-compulsory education and are home-owners. The low-SES individual has only compulsory education and is not a home-owner. The trend in the predicted mean of η_1 and η_2 across birth-cohorts for the low-SES man and woman is steep and upward. For the median-SES man and woman there is only a slightly upward trend. For high-SES woman the trend is almost flat; the trend for physical disability of the high-SES man is significantly downward.

Figure 3a: Predictions of the latent disability index by cohort of birth and SES for illustrative women aged 73



Notes: For definition of the illustrative individuals see text.

Figure 3b: Predictions of the latent disability index by cohort of birth and SES for illustrative men aged 73



Notes: For definition of the illustrative individuals see text.

4.3 The benefit receipt model

Table 11 reports structural parameters ($\gamma_1, \gamma_2, \boldsymbol{\beta_2}$) for the observed pattern of receipt of DBs (AA or DLA). They are semi-reduced form, representing both the take-up behaviour of disabled individuals and the decision-making of benefit claim assessment. They also include the possible reconsideration of the claim (by DWP or by external tribunals) in the case that the claim was unsuccessful at its initial attempt (Pudney, 2010) and possible errors in reporting behaviours.

The structural approach in use enables $\boldsymbol{\beta_2}$ to account only for the direct effect of $\boldsymbol{X}$ on DBs receipt, net of their indirect effects through their gradient with η_1 and η_2, captured by the coefficients $\boldsymbol{\beta_{1(.)}}$ in equation (1).

Receipt of DB is clearly disability-related. Physical disability consistently emerges as the dominant variable in explaining DBs receipt. The t-statistic associated with η_1 (η_2) is about 41 (11) for women and about 20 (9) for men. Controlling for latent disability, the estimated probability of receiving DBs increases with age but is significant (at the 1% level) only for women.

In line with findings documented elsewhere (Hancock *et al.*, 2015; Morciano *et al.*, 2015; Pudney, 2010), the probability of receiving DBs declines significantly by SES for women and men. Although AA/DLA are non-means-testing benefits, it seems that, controlling for disability status, SES-related differences in claim behaviour[76] make the AA/DLA programmes indirectly dependent on income.

For model A, receipt of DBs has a positive and similar birth-cohort trend for both women and men. For a given level of disability and holding other factors fixed, we found a cohort-year increases the probability of receiving DBs by about 0.3%. This direct effect[77] translates to an increase in the probability of receiving DBs of about 6.6% for the latest cohort of men and women born in 1945 relative to the one born in 1924. The total effect, estimated as the sum of the direct effect plus the indirect cohort-year effect on cognitive and physical disability on DBs receipt, is about 0.6% per cohort-year for women and 0.4% for men, corresponding to an increase in the probability of receiving DBs of about 13.2% and 8.8% for the cohort of women and men born in 1945 relative to the one born in 1924.

Based on the results of model B in Table 11, conclusions on the interaction effects are less straightforward to interpret than those for the latent disability equations, given the non-linearity of the probit function used for modelling DBs receipt (Ai & Norton, 2003) and the possible conflicting deductions that, even in linear models, can be drawn when using a F-test instead of a simpler t-test.

[76] Entitlement to AA/DLA should be independent of SES and therefore should not have a direct impact in the awarding process. The AA/DLA claim form does not explicitly require such information although they might be deducted and used by programme administrators. On the other hand, the likelihood of appealing against an unsuccessful claim might be positively related to SES.

[77] The term "direct effect" is meant to quantify the effect that is not mediated by other variables in the model including cohort-year effects in physical and cognitive disability.

Misleadingly, if we simply look at single coefficients associated with the interaction terms we will conclude that SES-cohort interactions are not significant for men. However, tests of no interaction effects[78] are rejected at the 1% level for all SES variables interacted also for men.

We estimate a negative, albeit negligible, coefficient for the interactions with income and housing wealth, implying attenuated birth-cohort trend in the receipt of DBs for homeowners with high income. On the other hand, the coefficient associated with the interaction with level of education is positive for women and men. This would suggest that successive cohorts of more educated individuals are more likely to be in receipt of DBs, *ceteris paribus*. However, the effect is partially counterbalanced by the negative relationship of level of education with DBs receipt.

[78] We carried out Wald tests for nonlinear models where the coefficients associated with birth-cohort, the SES variable and its interaction are tested to be zero.

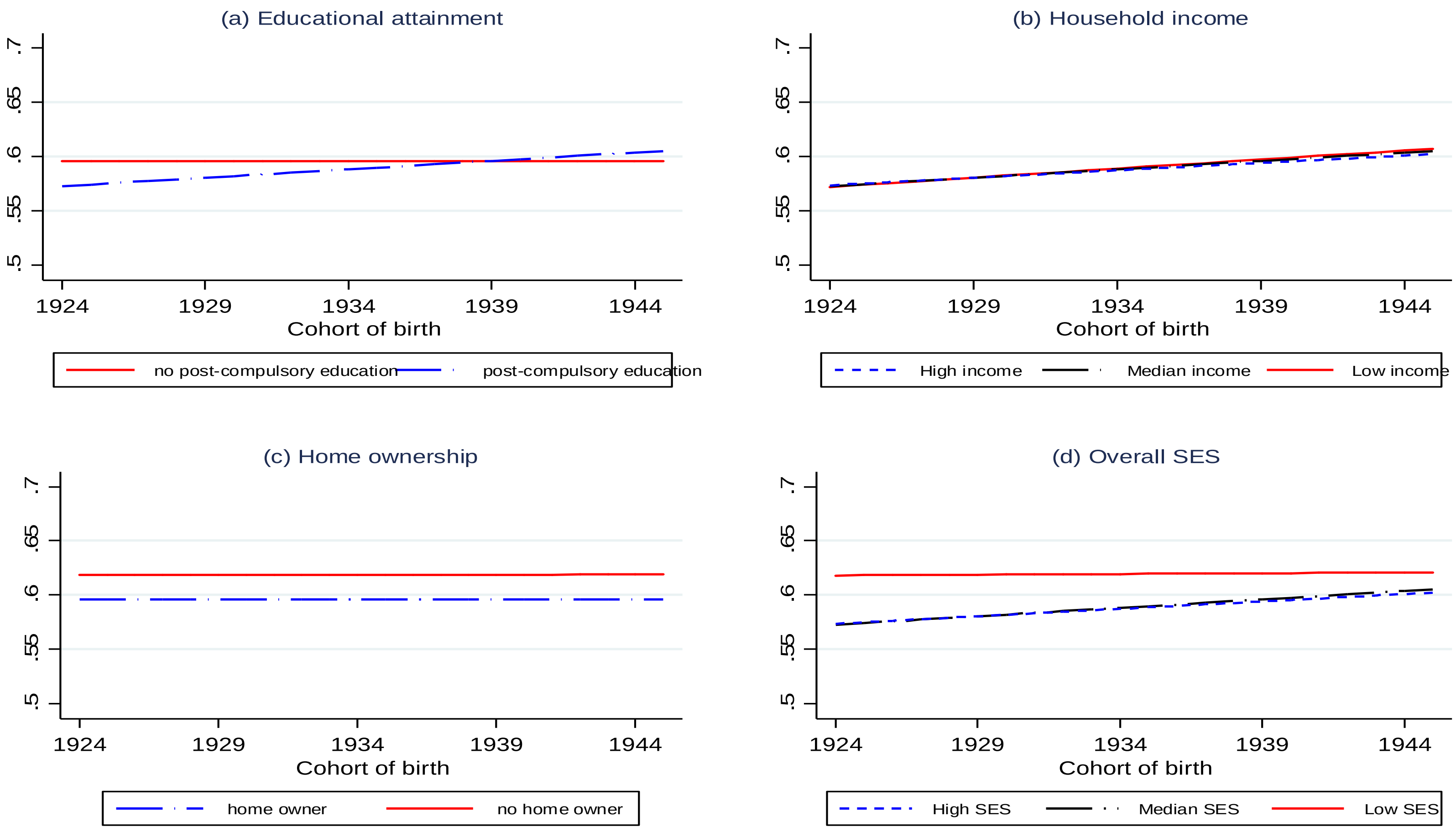
Table 11: Estimates from the DBs receipt equation

	Women		Men	
	Model A	Model B	Model A	Model B
Latent physical disability (η_1)	0.084***	0.084***	0.065***	0.066***
	(0.002)	(0.002)	(0.003)	(0.003)
Latent cognitive disability (η_2)	0.046***	0.046***	0.033***	0.033***
	(0.004)	(0.004)	(0.004)	(0.004)
Age	0.002***	0.002***	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Post-compulsory school	-0.031***	-0.064***	-0.023***	-0.034***
	(0.004)	(0.008)	(0.004)	(0.009)
Log household income[a]	-0.009**	0.007	-0.014***	-0.019**
	(0.003)	(0.007)	(0.003)	(0.007)
Home-ownership	-0.061***	-0.059***	-0.059***	-0.067***
	(0.004)	(0.009)	(0.005)	(0.010)
Cohabitation	-0.018***	-0.018***	-0.006	-0.005
	(0.003)	(0.003)	(0.004)	(0.004)
Scotland	0.030***	0.030***	0.030***	0.030***
	(0.004)	(0.004)	(0.004)	(0.004)
Wales	0.061***	0.061***	0.049***	0.049***
	(0.008)	(0.008)	(0.008)	(0.008)
Northern Ireland	0.132***	0.137***	0.101***	0.103***
	(0.007)	(0.007)	(0.007)	(0.008)
Birth-cohort	0.003***	0.007*	0.003***	-0.002
	(0.001)	(0.003)	(0.001)	(0.003)
Birth-cohort * post-compulsory school		0.004***		0.001
		(0.001)		(0.001)
Birth-cohort * income		-0.001*		0.001
		(0.000)		(0.001)
Birth-cohort * home-ownership		-0.000		0.001
		(0.001)		(0.001)
Constant	-0.499***	-0.809***	0.003	-0.245***
	(0.050)	(0.060)	(0.055)	(0.066)
$\hat{\sigma}^2_R$	0.106***	0.106***	0.092***	0.092***
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	52,229		44,504	

Notes: [a] For definition of household income see text. *Level of significance:* * p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors in parenthesis.

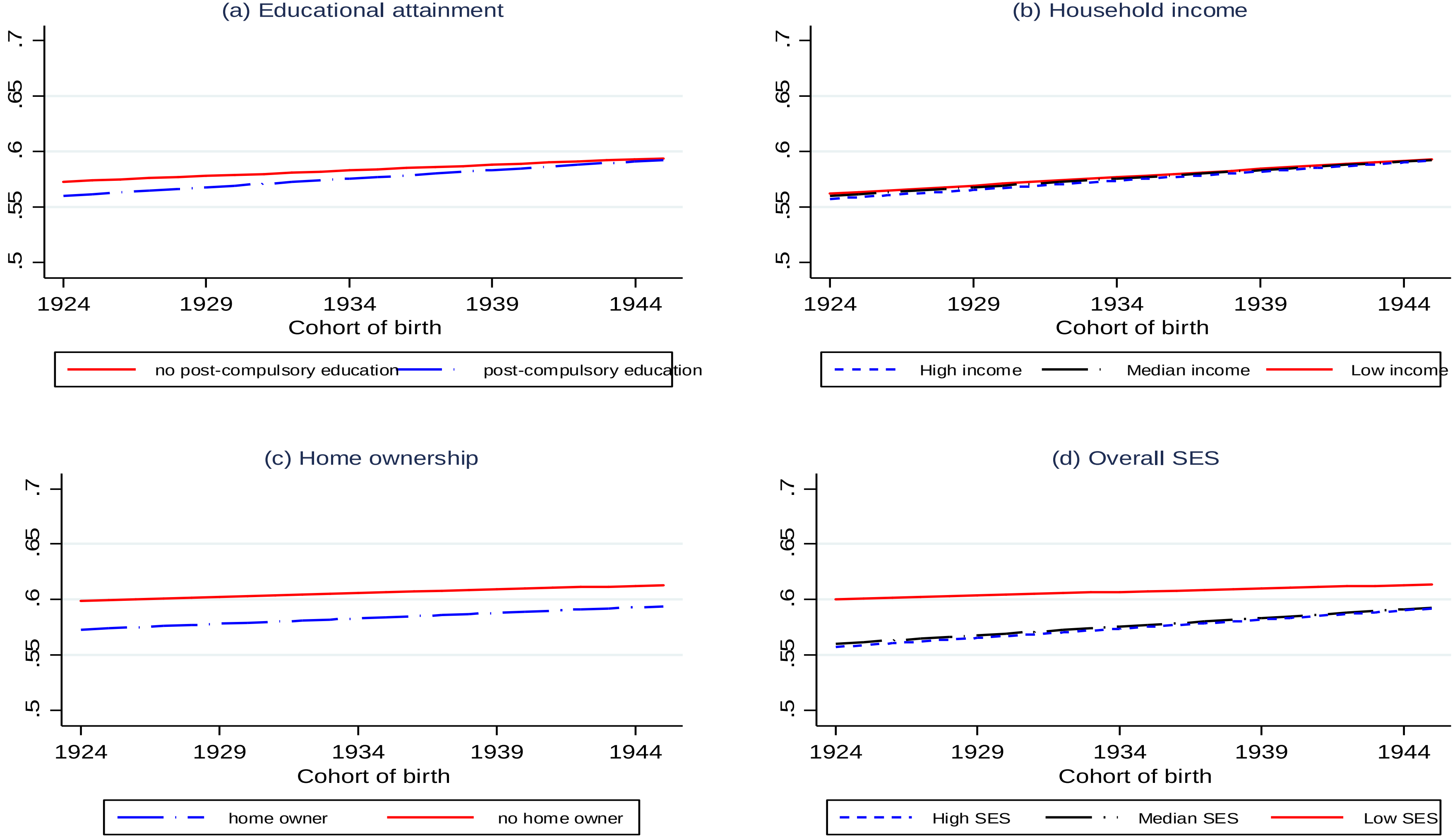
We make use of a graphical presentation to better assess the implications of specification B. Figure 4 highlights the main findings for the illustrative men and women mentioned above, assumed to have a level of physical and cognitive disability set at the 90$^{\text{th}}$ percentile of the disability indices ($\widehat{\eta_1}$, $\widehat{\eta_2}$) as predicted in the whole sample. First, at a given level of disability, the direct effect of SES on DBs receipt can be assessed from the SES-related differences in the predicted probabilities of receipt at any birth-cohort. Secondly, it is clear that the depicted positive birth-cohort trend in receipt of DBs differs only slightly according SES. For low SES, it is found almost flat for both women and men. On the other hand, a positive birth-cohort trend is found for median- and high-SES individuals, mainly thanks to a positive cohort-by-educational attainment effect in the receipt of DBs, in particular for women. For example, the ratio of the estimated probabilities of receipt for low-SES to median-SES for the cohort of women born in 1924 is 1.07 (.62/.57) which reduces to 1.03 (.62/.60) for the latest cohort. For men, we found a similar birth-cohort trend (see model A) but weaker evidence of diverging birth-cohort trends by SES. The ratio of the estimated probability of receipt of DBs for low-SES to median-SES for the cohort of men born in 1924 is only 1.08 (.57/.53); the same ratio computed for the latest cohort born in 1945 is 1.04 (.59/.57).

Figure 4a: Predictions of the probability of receiving DBs by cohort of birth and SES for illustrative highly disabled women aged 73



Notes: For definition of the illustrative individuals see text. η_1 and η_2 set at their 90$^{\text{th}}$ percentile of the gender-specific values.

Figure 4b: Predictions of the probability of receiving DBs by cohort of birth and SES for illustrative highly disabled men aged 73



Notes: For definition of the illustrative individuals see text. η_1 and η_2 set at their 90$^{\text{th}}$ percentile of the gender-specific values.

5 Summary and policy implications

We have analysed cohort-year effects in physical and cognitive disability and in the receipt of AA/DLA of older people born between 1924 and 1945 by pooling data from the Family Resource Surveys (FRS) carried out from 2002/3 to 2011/12.

The econometric approach in use incorporates a two-latent factor representation of the individual's disability in a system of structural equations. The individual's disability is characterised by correlated physical and cognitive dimensions that are measured by potentially error-contaminated self-reported FD indicators. Physical and cognitive disability, together with observable characteristics, determine receipt of DBs.

Our findings yield the following clear-cut messages that are relevant for current and planned policy reforms aimed at supporting older people with care needs.

Controlling for age and other relevant characteristics, we found evidence of a significant increase in physical and cognitive disability among the successive cohorts of older people in the UK. Increasing exposure to risk factors (e.g., obesity) and associated conditions might be the leading determinant. On the other hand, it is also possible that unfavourable conditions during infancy and childhood for the older cohorts had preselected the strongest members, explaining the observed increase for the younger compared with the older ones. It also should be noticed that prevalence of disability can also increase if the life expectancy of successive cohorts of people disabled earlier in life increases, even if the age of onset of disability is stable.

Whatever the reason underlying such a trend, the immediate policy implication that can be drawn is that it would not be prudent for policy-makers to count on future reductions in the prevalence of disability among elderly people to offset the rising demand for long-term care that will result from population ageing. It is also worrying to note that even a steady-state approach that projects the future number of disabled and the associated costs of disability programmes by using conditional rates observed at a single point in time could lead to severe underestimation. Previous projections of the public cost of long-term care in the UK have not taken cohort-year trends into account (e.g., Karlsson *et al.*, 2006; Pickard *et al.*, 2007; Wittenberg *et al.*, 2011).

The overall birth-cohort increase in disability hides a diverging gap between the socioeconomically advantaged and disadvantaged in later life. This is particularly evident for physical disability, especially for men: it increases among low SES individuals and decreases among high-SES individuals. This study provides no information about how such widening in SES differences in disability originated. Increasing exposure to unhealthy environments for low-SES individuals have been widely documented (Lynch *et al.*, 1997) but our findings might also reflect a reduction in mortality among low-SES disabled people. Additionally, our data cover only the private household population and we do not account for trends in the de-institutionalisation of care for older people. Some of the most severely disabled people live in care homes and there is evidence that some aspects of socio-economic advantage (e.g. home-ownership) reduce the risk of care home entry (Hancock *et al.*, 2002). If there were a substantial decrease in the proportion of the older population in care homes, it would partly explain the trends reported here. However, comparison of the 2001 and 2011 Census of the UK population shows that the (small) percentage of people over 65 resident in "medical and care" establishments fell only very slightly from 3.8% to 3.3%.[79] Even if all of this reduction consisted of low-SES individuals, it would explain only a very small part of the trends we find for the household population. Whatever the reason(s), if this widening trend continues it could have important implications for the future costs of the public system of care and support for people with care needs, since low-SES disabled individuals are more likely to be entitled to public support for the costs of their care.

The growth in AA/DLA receipt would not be inherently problematic if it were to reflect a rising incidence of physical and cognitive impairments or a lowering - among disabled - of the administrative and individual barriers associated with

[79] Calculated from 2001 and 2011 Census data available at:
http://www.scotlandscensus.gov.uk/ods-web/standard-outputs.html (Scotland);
http://www.nisra.gov.uk//Census/2001%20Census%20Results/StandardTables.html (Northern Ireland)
http://www.ons.gov.uk/ons/guide-method/census/2011/census-data/index.html (England and Wales).

their take-up. In both cases, in fact, target efficiency of the DB programmes is maintained or even improved.

We found that receipt of AA/DLA is strongly related to severity of (mainly physical) disability and increases with age. At old-age, long-term AA/DLA receipt often reflects the “absorbing state” of the underlying disabling conditions and the frailty associated with ageing. The message we draw is that the net effect on the public budget of reforms which introduce a regular re-assessment of the disabling condition at old age needs to be carefully designed, since the financial implication of such reforms crucially depends on the extent to which savings from the reduction in the *leakage* problems are able to offset the additional costs induced by the re-assessments.

Our econometric approach has allowed the separation of cohort-year and SES effects directly related to DBs receipt from the indirect ones that they would have via disability. We found that receipt of AA/DLA declines significantly by SES through a direct effect, due to take-up behaviours, and an indirect effect due to the SES-gradient with disability. Thus, reforms that include the options of making the benefit subject to income tax (Lloyd, 2014) or introduce means-testing for new claimants (Commission on the Future of Health and Social Care in England, 2014), while increasing administrative costs and stigma-related target inefficiencies, might have little impact on the financial sustainability of the system.

For a given level of disability and holding other factors fixed, we found a very small, albeit statistically significant, direct effect in the probability of receiving DBs of about 0.3% per cohort-year. It translates to an increase in the probability of receiving DBs of about 6.6% for the latest cohort of men and women born in 1945 relative to the one born in 1924. Accounting for the indirect cohort-year effect exerted thought disability, we estimate a total cohort-year effect in the probability of receiving DBs of about +13.2% and +8.8% for the cohort of women and men born in 1945 relative to the one born in 1924.

When allowing for SES-differential birth-cohort trends in DBs receipt, we find a statistically significant, albeit small, difference in the cohort-by-educational attainment effect and virtually no cohort-year changes by economic factors (income and home-ownership). This would suggest that later cohorts of individuals, at a given level of disability, might be taking more advantage of their level of education in navigating the disability benefits system or that they have lowered the perceived stigma from claiming benefits. In this view, the cohort-year effects we estimated might have had the desired effect of reducing inequality in the DBs take-up and thus improving target efficiency of the system.

References

Ai, C., & Norton, E. C. (2003). Interaction Terms in Logit and Probit Models. *Economics letters, 80* (1), 123-129.

Autor, D. H. (2011). *The Unsustainable Rise of the Disability Rolls in the United States: Causes, Consequences, and Policy Options.* (WP 17697). National Bureau of Economic Research

Bell, A., & Jones, K. (2013). The Impossibility of Separating Age, Period and Cohort Effects. *Social Science and Medicine, 93*, 163-165.

Berthoud, R. (2009). *Measuring the Impact of Disability Benefits: A Feasibility Study.* (2009-06). ISER Working Paper Series

Black, D., Daniel, K., & Sanders, S. (2002). The Impact of Economic Conditions on Participation in Disability Programs: Evidence from the Coal Boom and Bust. *The American Economic Review, 92* (**1**), 27-50.

Bland, J. M., & Altman, D. G. (1997). Statistics Notes: Cronbach's Alpha. *BMJ (Clinical Research Ed.), 314* (**7080**), 572.

Bound, J., & Burkhauser, R. V. (1999). Chapter 51 Economic Analysis of Transfer Programs Targeted on People with Disabilities. In C. A. Orley & C. David (Eds.), *Handbook of Labor Economics* (Vol. 3C, pp. 3417-3528): Elsevier.

Bruce, M. L., Seeman, T. E., Merrill, S. S., & Blazer, D. G. (1994). The Impact of Depressive Symptomatology on Physical Disability: Macarthur Studies of Successful Aging. *American Journal of Public Health, 84* (**11**), 1796-1799.

Burchardt, T. (1999). *The Evolution of Disability Benefits in the Uk: Re-Weighting the Basket.* CASEpaper (26). LSE STICERD

Burkhauser, R. V., & Daly, M. (2011). *The Declining Work and Welfare of People with Disabilities: What Went Wrong and a Strategy for Change*. Washington, D.C.: AEI Press.

Burkhauser, R. V., & Daly, M. C. (2012). Social Security Disability Insurance: Time for Fundamental Change. *Journal of Policy Analysis and Management, 31* (**2**), 454-461.

Burkhauser, R. V., Daly, M. C., McVicar, D., & Wilkins, R. (2014). Disability Benefit Growth and Disability Reform in the Us: Lessons from Other Oecd Nations. *IZA Journal of Labor Policy, 3* (**1**), 1-30.

Cleveland, W. S. (1979). Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association, 74* (**368**), 829-836.

Commission on the Future of Health and Social Care in England. (2014). *A New Settlement for Health and Social Care. Interim Report.* King's Fund

Crimmins, E. M., & Beltrán-Sánchez, H. (2011). Mortality and Morbidity Trends: Is There Compression of Morbidity? *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 66B* (**1**), 75-86.

Crimmins, E. M., Hayward, M. D., Hagedorn, A., Saito, Y., & Brouard, N. (2009). Change in Disability-Free Life Expectancy for Americans 70-Years-Old and Older. *Demography, 46* (**3**), 627-646.

Crimmins, E. M., Kim, J. K., & Solé-Auró, A. (2011). Gender Differences in Health: Results from Share, Elsa and Hrs. *The European Journal of Public Health, 21* (**1**), 81-91.

d'Uva, T. B., Lindeboom, M., O'Donnell, O., & Van Doorslaer, E. (2011). Slipping Anchor? Testing the Vignettes Approach to Identification and Correction of Reporting Heterogeneity. *Journal of Human Resources, 46* (**4**), 875-906.

Deaton, A. (2002). Policy Implications of the Gradient of Health and Wealth. *Health Affairs, 21* (**2**), 13-30.

Department for Work and Pensions. (2013a). Frs Response Rates. In frs_2011_12_introduction_family_resources_survey_june_13.pdf (Ed.), (pp. FRS Response Rates). Introduction to the Family Resources Survey: Department for Work and Pensions.

Department for Work and Pensions. (2013b). *Households Below Average Income: An Analysis of the Income Distribution 1994/95 -2011/12*. London: Department for Work and Pensions.

Department for Work and Pensions. (various years). Family Resources Survey. from Department for Work and Pensions http://discover.ukdataservice.ac.uk/series/?sn=200017

Falkingham, J., Evandrou, M., McGowan, T., Bell, D., & Bowes, A. (2010). *Demographic Issues, Projections and Trends: Older People with High Support Needs in the Uk*. Report for the Joseph Rowntre Foundation: ESRC Centre for Population Change.

Feinstein, J. S. (1993). The Relationship between Socioeconomic Status and Health: A Review of the Literature. *Milbank Quarterly, 71* (**2**), 279-322.

Fitzgerald, J. F., Smith, D. M., Martin, D. K., Freedman, J. A., & Wolinsky, F. D. (1993). Replication of the Multidimensionality of Activities of Daily Living. *Journal of Gerontology, 48* (**1**), S28-S32.

Freedman, V. A., Martin, L. G., & Schoeni, R. F. (2002). Recent Trends in Disability and Functioning among Older Adults in the United States: A Systematic Review. *JAMA, 288* (**24**), 3137-3146.

Glenn, N. D. (1976). Cohort Analysts' Futile Quest: Statistical Attempts to Separate Age, Period and Cohort Effects. *American Sociological Review, 41* (**5**), 900-904.

Hancock, R., Arthur, A., Jagger, C., & Matthews, R. (2002). The Effect of Older People's Economic Resources on Care Home Entry under the United Kingdom's Long-Term Care Financing System. *The journals of gerontology. Series B, Psychological sciences and social sciences, 57* (**5**), S285-S293.

Hancock, R., Morciano, M., Pudney, S., & Zantomio, F. (2015). Do Household Surveys Give a Coherent View of Disability Benefit Targeting? A Multi-Survey Latent Variable Analysis for the Older Population in Great Britain. *Journal of the Royal Statistical Society. Series A, Statistics in Society, DOI: 10.1111/rssa.12107.*

Hancock, R., & Pudney, S. (2014). Assessing the Distributional Impact of Reforms to Disability Benefits for Older People in the Uk: Implications of Alternative Measures of Income and Disability Costs. *Ageing and Society, 34*, 232-257.

Haveman, R. H., & Wolfe, B. L. (1984). The Decline in Male Labor Force Participation: Comment. *Journal of Political Economy, 92* (**3**), 532-541.

HMSO. (1988). *Benefit for Disabled People: A Strategy for Change.* Her Majestry's Stationery Office

Jarvis, C., & Tinker, A. (1999). Trends in Old Age Morbidity and Disability in Britain. *Ageing and Society, 19* (**5**), 603-627.

Johnson, R. J., & Wolinsky, F. D. (1993). The Structure of Health Status among Older Adults: Disease, Disability, Functional Limitation, and Perceived Health. *Journal of Health and Social Behavior*, 105-121.

Juhn, C. (1992). Decline of Male Labor Market Participation: The Role of Declining Market Opportunities. *The Quarterly Journal of Economics, 107* (**1**), 79-121.

Karlsson, M., Mayhew, L., Plumb, R., & Rickayzen, B. (2006). Future Costs for Long-Term Care: Cost Projections for Long-Term Care for Older People in the United Kingdom. *Health Policy, 75* (**2**), 187-213.

Kasparova, D., Marsh, A., & Wilkinson, D. (2007). *The Take-up Rate of Disability Living Allowance and Attendance Allowance: Feasibility Study (Research Report No 442).* London: Department for Work and Pensions.

King, G., Murray, C. J., Salomon, J. A., & Tandon, A. (2004). Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research. *American Political Science Review, 98* (**01**), 191-207.

Lloyd, J. (2014). *Options for Funding Care.* Paper for Commission on the Future of Health and Social Care in England. The King's Fund

Lynch, J. W., Kaplan, G. A., & Salonen, J. T. (1997). Why Do Poor People Behave Poorly? Variation in Adult Health Behaviours and Psychosocial Characteristics by Stages of the Socioeconomic Lifecourse. *Social Science and Medicine, 44* (**6**), 809-819.

Martin, L. G., Freedman, V. A., Schoeni, R. F., & Andreski, P. M. (2010). Trends in Disability and Related Chronic Conditions among People Ages Fifty to Sixty-Four. *Health Affairs, 29* (**4**), 725-731.

Martin, L. G., Schoeni, R. F., Andreski, P. M., & Jagger, C. (2012). Trends and Inequalities in Late-Life Health and Functioning in England. *Journal of Epidemiology and Community Health, 66* (**10**), 874-880.

McVicar, D. (2008). Why Have Uk Disability Benefit Rolls Grown So Much? *Journal of Economic Surveys, 22* (**1**), 114-139.

Morciano, M., Hancock, R., & Pudney, S. (2015). Disability Costs and Equivalence Scales in the Older Population in Great Britain. *Review of Income and Wealth, 61* (**3**), 494-514.

Morciano, M., Hancock, R. M., & Pudney, S. E. (2015). Birth-Cohort Trends in Older-Age Functional Disability and Their Relationship with Socio-Economic Status: Evidence from a Pooling of Repeated Cross-Sectional Population-Based Studies for the Uk. *Social Science and Medicine, 136-137C* (available at Earlyview).

Nunnally, J. C., Bernstein, I. H., & Berge, J. M. t. (1967). *Psychometric Theory* (Vol. 226): McGraw-Hill New York.

Oksuzyan, A., Crimmins, E., Saito, Y., O'Rand, A., Vaupel, J. W., & Christensen, K. (2010). Cross-National Comparison of Sex Differences in Health and Mortality in Denmark, Japan and the Us. *European Journal of Epidemiology, 25* (**7**), 471-480.

Parsons, D. O. (1980). The Decline in Male Labor Force Participation. *Journal of Political Economy, 88* (**1**), 117-134.

Pattison, D., & Waldron, H. (2013). Growth in New Disabled-Worker Entitlements 1970-2008. *Soc. Sec. Bull., 73* (**4**), 25-48.

Pickard, L., Comas-Herrera, A., Costa-Font, J., Gori, C., di Maio, A., Patxot, C., Wittenberg, R. (2007). Modelling an Entitlement to Long-Term Care Services for Older People in Europe: Projections for Long-Term Care Expenditure to 2050. *Journal of European Social Policy, 17* (**1**), 33-48.

Pudney, S. (2010). *Disability Benefits for Older People: How Does the Uk Attendance Allowance System Really Work?* (2010-02). ISER Working Paper Series

Schoeni, R. F., Freedman, V. A., & Wallace, R. B. (2001). Persistent, Consistent, Widespread, and Robust? Another Look at Recent Trends in Old-Age Disability. *The journals of gerontology. Series B, Psychological sciences and social sciences, 56* (**4**), S206-S218.

Schoeni, R. F., Martin, L. G., Andreski, P. M., & Freedman, V. A. (2005). Persistent and Growing Socioeconomic Disparities in Disability among the Elderly: 1982–2002. *American Journal of Public Health, 95* (**11**), 2065-2070.

Spector, W. D., & Fleishman, J. A. (1998). Combining Activities of Daily Living with Instrumental Activities of Daily Living to Measure Functional Disability. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 53B* (**1**), S46-S57.

Stevens, J. P. (2009). *Applied Multivariate Statistics for the Social Sciences* (5th ed.). New York: Routledge.

Stoop, I. A. L. (2005). *The Hunt for the Last Respondent: Nonresponse in Sample Surveys*. The Haque: SCP, Social and Cultural Planning Office of the Netherlands.

Verbrugge, L. M., & Jette, A. M. (1994). The Disablement Process. *Social Science and Medicine, 38* (**1**), 1-14.

Wanless, D. (2006). *Securing Good Care for Older People: Taking a Long-Term View.* King's Fund

WHO. (2011). *World Report on Disability*. Geneva, Switzerland: World Health Organization.

WHO. (2014). Review of Social Determinants and the Health Divide in the Who European Region: Final Report. Copenhagen: WHO Regional Office for Europe.

Wittenberg, R., Hu, B., Hancock, R., Morciano, M., Comas-Herrera, A., Malley, J., & King, D. (2011) Projections of Demand for and Costs of Social Care for Older People in England, 2010 to 2030, under Current and Alternative Funding Systems. *PSSRU discussion paper, 2811/2*. PSSRU, London, UK.

Zaninotto, P., Nazroo, J., & Banks, J. (2010). 7. Trends in Disability. In J. Banks, C. Lessof, J. Nazroo, N. Rogers, M. Stafford & A. Steptoe (Eds.), *Financial Circumstances, Health and Well-Being of the Older Population in England* (pp. 254-274). London: The Institute for Fiscal Studies.

Appendix

Figure A1: The path diagram of the two-latent factor model

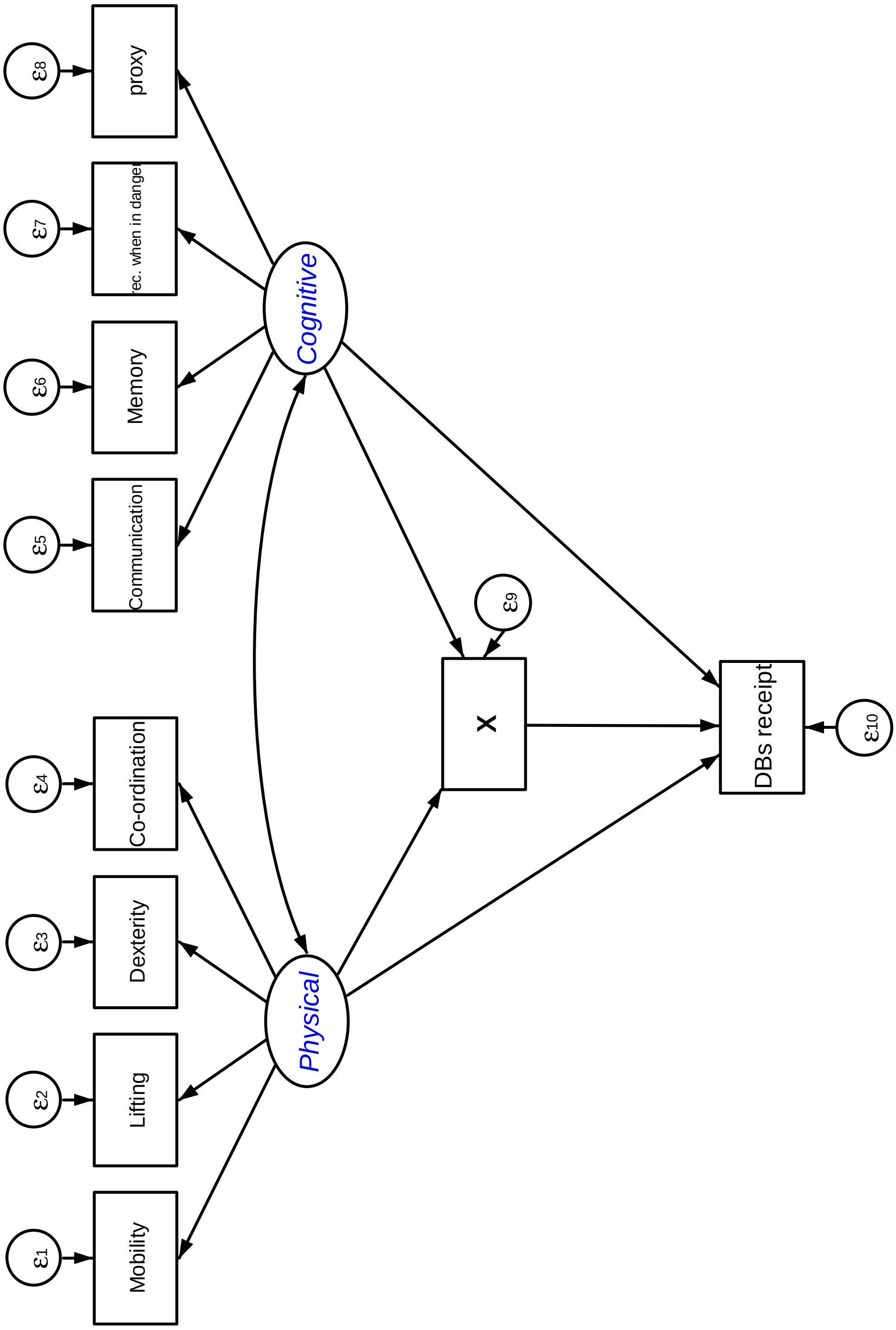
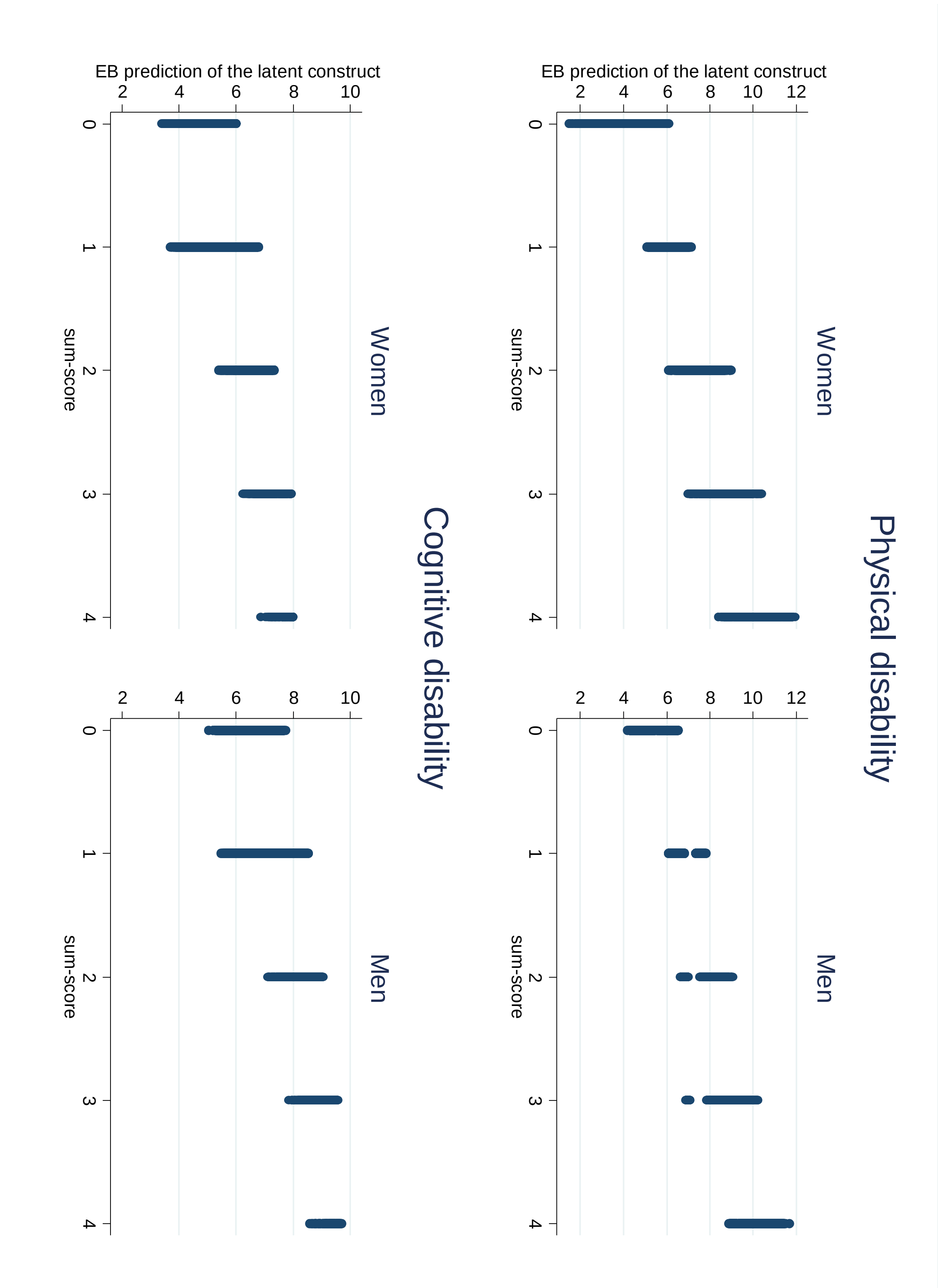


Figure A2: Correlation between Empirical Bayes and sum-score disability scale (two-latent factor)



Conclusions and implications for future research

There is a pressing need for robust evidence to inform the policy discussion around the public system of care and support for older people with care needs.

One of the difficulties faced by researchers is that concepts like "disability" and "well-being" involved in this kind of research cannot be observed directly. Instead researchers must draw inferences on them by using a battery of imperfect survey indicators. The measurement noise in these indicators could cause bias in analytical results and lead to distorted policy recommendation, if the appropriate statistical methods are not used.

This thesis presented four empirical studies applied to the economics of disability in old age that make use of Structural Equation Models (SEM) with latent variables. Commonly employed in psychology and the social sciences, this approach is becoming increasingly important in economics, although it is still underused in health-related studies (Wang & Wang, 2012).

From the applied point of view, a latent factor SEM recognises that important theoretical concepts (e.g. disability) are latent rather than directly observable. By defining structural relations between indicators and embedding those latent concepts in a set of simultaneous equations, this approach is able to capture multiple latent variables simultaneously (Chapters 1 and 3) and operationalise a multi-dimensional notion of disability (Chapters 2 and 4). It can also handle difficult data with a sparse and noisy set of observed indicators (Chapters 1 and 2); allow the testing of hypotheses on the correlation among latent factors (Chapter 1); provide a basis for testing parameter invariance across different surveys (chapter 2) and population sub-groups (Chapter 3); and generate estimates of direct, indirect and total effects (Chapter 4). The traditional reluctance of economists to use self-reported subjective information can be overcome by allowing for the presence of measurement errors in the set of indicators used to operationalise latent concepts and by accounting for different respondents' behaviour in the self-evaluation activity demanded in a survey.

There are, however, some important drawbacks to this approach. One could argue that it is usually hard to interpret latent indexes because they have no natural scale. In all chapters we have shown that this problem can be handled naturally by ranking sample members according to their model-based posterior prediction of the latent constructs. Going beyond classical "data-driven" statistical techniques (e.g. principal component analysis), such indexes can supersede or complement more common weighted count measures that use equal, arbitrary or expert opinion weights (Decancq & Lugo, 2013).

In the empirical applications presented here the indicators of the underlying latent variables were generally found to be strongly correlated and to capture plausibly a single latent construct. However, there were some cases of indicators that were poorly correlated with the hypothesised construct, with low factor loadings and large error variances. This can signal poor indicator quality, but it can also be a symptom of underspecifying the number of latent factors: there is multi-dimensionality of the latent concept, and the dimensions are not perfectly correlated. Chapters 2 and 4 show the relevance of this issue in conceptualising the dimensionality of disability: allowing for two (rather highly correlated) dimensions plausibly interpreted as physical and cognitive disability improved the explanatory power of the model and our understanding of the underlying process.

A great deal of influential health research has been based on survey data and the increase of drop-out rates among older people in longitudinal studies, albeit mostly neglected by applied researchers, is a matter of concern.

The analysis of survey participation proposed in Chapter 1 has shown how understanding the process of "being surveyed" is not straightforward but offers a way of dealing with panel attrition. It highlights the importance of considering carefully individuals' attitudes and beliefs towards survey participation and the relationship of those attitudes with the outcomes of interest. But attitudes are themselves theoretical concepts which are only imperfectly measured by available indicators. As far as the econometric framework is concerned, this is one of the first applications of a latent factor SEM to the study of survey participation and it should be interpreted as a call for further research. Results from Chapter 1 have immediate implications for future applied research. We have showed that the effect of non-response is to bias downwards estimates of the prevalence of disability and receipt of related benefits and to attenuate the estimated socio-economic gradient in health. Handling (or at least recognising) this potential source of bias is of paramount importance in applied research aiming to formulate policy recommendations. Contact with the ELSA survey designers has confirmed their willingness to consider ways of retaining participants and enhancing the weighting adjustment procedure, by building upon this research.[80] However, whether new weights would help in other research areas is a matter for speculation. Future research should assess how generalisable our conclusions on sample selection bias are in static and dynamic analyses on mental health, life satisfaction, well-being and other outcomes highly correlated with respondents' engagement.[81]

[80] The collaboration would also enable the use of information not currently available to researchers which, by being good predictors of future non-response (e.g. number of calls before arranging an interview, interview length, interviewers' characteristics and their perception on respondents' level of engagement), would be valuable instruments for informing post-collection adjustment procedures for non-response. Derived weights from this chapter will shortly be made publicly available through the UK Data Archive.

[81] The analysis can make use of many surveys given that psychometric indicators and data on the respondents' level of engagement are now collected and publicly available for many multi-purpose surveys. As an example, it would interesting to assess the relevance of sample selection problems in the UK component of the European Union Statistics on Income and Living Conditions (EU-SILC) following the decision of selecting SILC new sample units from the FRS.

Chapter 2 showed that, despite the considerable differences between three major surveys in the number of disability questions they use, the wording of those questions, the way they select their samples and the way they handle cases where the subject is unable to answer personally, the results give a similar statistical picture of the relationship between disability and receipt of the main disability benefit received at old-age, the Attendance Allowance. This robustness is a very encouraging finding for policy analysts, given the proliferation of disability scales (mainly in the clinical epidemiologic literature), the cost of designing new questionnaires and of collecting reliable information through surveys; and the "*almost irresistible pressure* [on politicians] *to cherry-pick – or even misrepresent – evidence*"[82].

The statistical approach proposed in Chapter 2 could easily be extended in other research domains, dealing with situations where conflicting results emerge from indicators available in different surveys, but also when changes in the questionnaire occurring from one sweep to another of the same survey prevents direct identification of changes of the same underlying phenomenon through time. This is particularly relevant to work based on the Family Resources Survey, which has recently changed the design of questions on functional difficulties and standard of living.

[82] Hancock, R., Morciano, M., Pudney, S., & Zantomio, F. (2013). "*Is cherry-picking disability data at all fruitful?"*, Society Central, https://societycentral.ac.uk/2013/07/04/is-cherry-picking-disability-data-at-all-fruitful/.

Chapter 3 showed how statistical well-being models can be used to estimate the extra costs of normal functioning associated with disability.

The clear result is that current public provision of cash disability benefits falls considerably short of total disability costs for older disabled people in Great Britain. Future research could make use of these estimates to assess the targeting and redistributive efficiencies of existing disability-related public programmes and ways in which to improve their policy design.

From an econometric point of view it would be interesting to extend the estimation procedure of the extra costs of disability in two dimensions. A panel dimension of the data could allow for unobserved individual heterogeneity (see e.g., Cullinan *et al.*, 2011) and permit a better understanding of the dynamics of disability and the process of adaptation (Easterlin, 1974) or "physical condition-neglect" (Sen, 1985, p.21) of standards of living.

Together with measures of individual happiness and life satisfaction, a survey such as the ELSA collects two types of measure to assess household welfare: the budget share for three types of good (income spent on food, clothes and leisure), and deprivation indicators similar to the FRS ones.[83] Such measures, while capturing different (but related) concepts of "well-being" (see e.g., Decancq *et al.*,

[83] As an example: "*please say how often you find you have too little money to spend on: First choices of food items; Have family and friends round for a drink or meal; Have an outfit to wear for social or family occasions; Keep your home in a reasonable state of decoration; Replace or repair broken electrical goods; Pay for fares or other transport costs to get to and from places you want to go; Buy presents for friends or*

2015; Pudney, 2011) are likely to provide very different estimates of valuations obtained through the compensating or equivalent variation principle. A direction of research is likely to discuss the "existence" and the "meaning" of such valuations, which can be empirically grounded in the ELSA data systematically.

Chapter 4 documented the relevance of birth-cohort trends in physical and cognitive functionings and in the receipt of non-means-tested cash disability benefits in old age in the UK. It shows the existence of diverging trends of functional disability by socio-economic status, with a steep increase in physical and cognitive disability among the disadvantaged. If such trends are likely to persist in the future, they would have tremendous implications for future costs of public programmes aimed at supporting people with care needs.

Projecting the economic implications of reforms of the system of care and support requires "realistic" assumptions on the evolution of disability for the coming decades.[84] The setup used in this chapter could be a promising way forward to build a dynamic microsimulation model that, while being better embedded in existing theoretical frameworks, could provide projections under a wide range of scenarios.

---

family once a year; Take the sorts of holidays you want; Treat yourself from time to time", with potential responses being: *Never, Rarely, Sometimes, Often* and *Most of the time*.

[84] "*Making projections into the future [...] requires a theory of how things unfold [...]. These should draw on the latest and best research. Moreover, the causal stories have to be empirically grounded and represented quantitatively [...]. Members of the research community have to be engaged.*" (Harding & Gupta, 2007).

Chapter 4 also provides evidence that, controlling for disability, the probability of receiving cash disability benefits has increased only slightly, with a more favourable trend for the better educated and virtually no cohort-year changes by economic factors (income and home-ownership). This might indicate differences in claiming behaviour, in assessment criteria, or in the way receipt is reported in a survey. Further research should explore these aspect further, for example using FRS data linked with Department of Work and Pensions (DWP) administrative data.[85]

Finally, the current policy discussion on whether and how to integrate disability benefits (currently administered by the DWP) with the Local Authority-administered system of social care (Barker, 2014) should be fed with analyses. It has often been suggested that, in comparison with social care, disability benefits are not well targeted to those disabled and in most financial need (see e.g., Department of Health, 2009, 2013; Wanless, 2006) in the light of supposed better target efficiency achieved by LA-subsidised care services. The new information on receipt of social care collected within the HSE and the ELSA would enable the *joint* evaluation of the target efficiency of the two programmes and might be used to inform the public debate.

[85] A linked dataset has recently been provided by the Department for Work and Pensions.

References

Barker, K. (2014). *A New Settlement for Health and Social Care.* King's Fund, London

Cullinan, J., Gannon, B., & Lyons, S. (2011). Estimating the Extra Cost of Living for People with Disabilities. *Health Economics, 20* (**5**), 582-599.

Decancq, K., Fleurbaey, M., & Schokkaert, E. (2015). Chapter 2 - Inequality, Income, and Well-Being. In B. A. Anthony & B. François (Eds.), *Handbook of Income Distribution* (Vol. Volume 2, pp. 67-140): Elsevier.

Decancq, K., & Lugo, M. A. (2013). Weights in Multidimensional Indices of Wellbeing: An Overview. *Econometric Reviews, 32* (**1**), 7-34.

Department of Health. (2009). *Shaping the Future of Care Together.* The Stationery Office, London

Department of Health. (2013). *Policy Statement on Care and Support Funding Reform and Legislative Requirements.* The Stationery Office, London

Easterlin, R. A. (1974). Does Economic Growth Improve the Human Lot? Some Empirical Evidence. In P. David & M. W. Reder (Eds.), *Nations and Households in Economic Growth:Essays in Honor of Moses Abramovitz* (pp. 89–125). New York: Academic Press.

Harding, A., & Gupta, A. (2007). *Modelling Our Future: Population Ageing, Social Security and Taxation*: Elsevier Science.

Pudney, S. (2011). Perception and Retrospection: The Dynamic Consistency of Responses to Survey Questions on Wellbeing. *Journal of Public Economics, 95* (**3–4**), 300-310.

Sen, A. K. (1985). *Commodities and Capabilities*. New York: North Holland.

Wang, J., & Wang, X. (2012). *Structural Equation Modeling: Applications Using Mplus*: John Wiley & Sons.

Wanless, D. (2006). *Securing Good Care for Older People: Taking a Long-Term View.* King's Fund, London